Beliefs and Uncertainty in Stochastic Modeling

by

Alexander Jeno Munk

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
2017

Doctoral Committee:

Professor Erhan Bayraktar, Chair
Professor Romuald S. Elie, Université Paris-Est Marne-la-Vallée
Assistant Professor Sergey Nadtochiy
Associate Professor Stilian A. Stoev
Professor Virginia R. Young

Alexander J. Munk
amunk@umich.edu
ORCID iD: 0000-0003-1328-0649

To my family

ACKNOWLEDGEMENTS

Studying with Professor Bayraktar has been an extraordinary journey. His curiosity was a continual source of inspiration and always drove me to tackle the next challenge in our research. He supported my learning and career progression at every stage and in every sense, helping me to grow in countless ways. He never stopped encouraging me to aim higher and push farther, which even by itself, had a seismic impact on our final results. He gave me the freedom to both try and fail, giving me resilience for the odyssey ahead. Perhaps most important, though, it was simply a joy to work with him. I am truly fortunate to have had such a remarkable advisor.

I am sincerely grateful to my reader Professor Romuald Elie, as well as my other committee members Professors Sergey Nadtochiy, Stilian Stoev, and Virginia Young, for their insightful feedback on my thesis. Their questions led me to interpret and understand my findings in entirely new ways. My perspective on potential future directions for the work was also markedly improved by their suggestions.

Being able to collaborate with University of Michigan's vibrant Financial and Actuarial Mathematics community including Bahman Angoshtari, Asaf Cohen, Professor Joseph Conlon, Sebastian Herrmann, Professor Mattias Jonsson, Christian Keller, Jiaqi Li, Professor Kristen Moore, Professor Johannes Muhle-Karbe, Jinniao Qiu, Yavor Stoev, Gu Wang, Xiang Yu, Yuchong Zhang, and Zhou Zhou was invigorating. Their seminars have granted me the opportunity to gain exposure to a number of thrilling discoveries in financial mathematics. Our subsequent discussions, which were invariably filled with their subtle observations, have substantially expanded my grasp of our field.

The financial support afforded by my Rackham Predoctoral Fellowship, Mathematics Department Graduate Fellowship, and advisor dramatically facilitated my progress in the doctoral program. It permitted me to fully immerse myself in research throughout my time in graduate school. Being able to so wholly focus on my thesis developed my perspective in a way that would have been unattainable for me, otherwise.

I would like to express my appreciation for the staff of the Department of Mathematics at the University of Michigan, particularly Tara McQueen, Brenae Smith, Molly Tiernan, and Beste Windes. Their never-ending assistance considerably simplified the navigation of the doctoral program's logistical aspects. Their advice regarding key campus events and resources proved invaluable in my career search as well.

My colleagues during my summer internship at Goldman Sachs, particularly Seung-jae Bang, Xin Chen, Jaehyung Choi, Cecilia Feng, Hamed Firouzi, Stephane Floret, Sherry Hu, Zhitao Lu, Murad Nayal, Piyush Sachdeva, Ryan Sankarpersad, Bode Sule, and Yijie Zhou, transformed my understanding of financial mathematics in practice. It was a boon to be the beneficiary of their patient mentorship. Their insightful guidance deeply altered my perspective on how to model real-world phenomena and profoundly affected my final thesis project.

I have thoroughly enjoyed being the long-time officemate of Dondi Ellis and Roman Gayduk. Over the years, I have considerably benefited from our discussions of course material, approaches to various aspects of the doctoral program, and career strategies. It has also been a genuine pleasure to enjoy their friendship.

I am grateful to my future colleagues at Chicago Trading Company including Jessica De Maria, Erik Doeff, Meghan Fore, Jeffrey Jones, Anurag Kumar, Andrew Marantan, Serge Posudevsky, Gregory Sobczak, and Adam Westergaard. Graduate school has been a tremendous adventure, but knowing that I will soon have the opportunity to continue building on my studies and be a part of such a wonderful team has made it even more exciting. I am looking forward to joining you all soon.

I am obliged to Professors Caroline Klivans, Raghavan Narasimhan, and Madhav Nori who were instrumental in my transition to graduate school. It was transformative to be engaged by their erudite lectures at the University of Chicago. Later, when they supported my application to the University of Michigan, they continued to markedly impact my educational path.

I learned a great deal from my former colleagues at Harlem Academy and would especially like to thank Chris Cunningham, Vinny Dotoli, Allison Duarte, Sarah Lefeber, Jeffrey Newman, Sean Robertson, and Whitney Wood. In addition to research, teaching is a critical part of the doctoral program experience. Trying to incorporate all that I could from their inspired, committed examples helped me and my students at the University of Michigan significantly.

I want to acknowledge Professor H. Bernhard Schlegel and his former research group at Wayne State University, notably Professors Smriti Anand, Hrant Hratchian, and Xiaosong Li. They generously devoted their time to enable me to transition from a high school introductory course to computational chemistry research. In the process, they shared their passions for tackling seemingly impossible challenges and contributing to our body of knowledge.

Professor Paul Sally, my undergraduate advisor at the University of Chicago, was simply a one-of-a-kind force of nature. I have never met someone who loved both

mathematics and education with such fiery intensity. He was selfless and would do anything in his power (or even beyond) to help his students, Chicago Public Schools, and the broader mathematics community. I could not overestimate his influence on me and am truly indebted to him.

Most of all, I would like express my thanks to Michèle, Dad, Mom, Byedaddy, Gaga, Grandad, Gramme, Uncle Saul, Aunt Robin, Jean, Louise, and Élyane. You ignited my zest for math long ago by *not* letting me win as we played quantitatively-inclined children's games. You arranged for countless enrichment activities as I grew up, which cemented my need to continuously learn. You endlessly offered care and comfort, which made it possible for me to follow my curiosity wherever it led. You provided wise counsel at every critical juncture, which substantively improved my trajectory. You patiently understood when I spent our vacations working on my latest project, which allowed me to complete my Ph.D. I am lucky to be a part of your lives and dedicate this thesis to you.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

CHAPTER I

Introduction

Belief specification, as well as the identification of sources and statistical properties of uncertainty, is a crucial stage in stochastic model development. In much of the classical literature, one would begin by pinpointing a future event whose outcome would effectively determine the conclusion of the scenario of interest. It would then be necessary to hypothesize a particular distribution for the event's outcomes. Once these decisions had been made, answering questions with both theoretical and practical significance became a matter of careful argumentation and computation.

In this thesis, we are concerned with the following two questions, especially in cases when they can be motivated by financial applications: What if one is unable to select a single distribution which most appropriately characterizes the likelihoods of future outcomes? What if one has made a choice, even the best conceivably available choice, but it is simply wrong? For example, the first issue could naturally arise when several distribution candidates appear equally plausible, but they all have unique advantages and flaws. The occurrence of the second issue is nearly inevitable: As noted by the legendary statistician George Box, "All models are wrong" ([59]).

The financial mathematics community has investigated our first question since the seminal works of Avellaneda et al. ([28]) and Lyons ([165]). Volatility is a key parameter when pricing certain securities including options, and these papers examine what unfolds, if the volatility is not precisely known. Roughly, an agent faced with this problem is driven to consider several possible distributions for his future event's outcomes. He considers infinitely many of them, in fact, and some may badly disagree with others. For instance, one of his distributions may deem the occurrence of a particular outcome as certain, while another views it as impossible. The financial implications of such a framework and its variants have been thoroughly studied ([18], [42], [52], [53], [91], [109], [125], [126], [172], [203], and [124]). In support of this line of research and due to its independent relevance, others have investigated the transference of statistical properties from classical objects to their counterparts under this new scheme ([190], [191], [223], [224], [195], [241], [181], [127], [94]). Chapters II and III fall into the latter category. Intuitively, they primarily focus on features of large aggregations of future events where the corresponding distributions are uncertain.

To the best of our knowledge, our second question has attracted less scholarly attention in the contexts of index tracking during a reconstitution, parimutuel wagering, or mini-flash crashes. Briefly, index funds like the Vanguard 500 Index Fund (VFINX) aim to replicate a chosen market benchmark such as the S&P 500 ([230]). Parimutuel wagering is a popular betting system used in finance ([38]), sports ([7]), lotteries ([6]), and prediction markets ([197]). Mini-flash crashes are violent, rapid spikes or crashes in the price of some security, e.g., the 300ms swing in Qualys, Inc. (QLYS) from \$10 to \$0.0001 and back, which took place on April 25, 2013 ([9]). Chapters IV - VI can be viewed as suggesting that it is natural to make mistakes in these situations, whether by picking a seemingly reasonable (but imperfect) objective or relying upon a sophisticated (but faulty) model. These innocuous errors can have surprising and occasionally disastrous consequences.

We conclude Chapter I with a more precise outline of the remainder of the thesis.

1.1 Chapter II Summary

In one dimension, the theory of the G-normal distribution is well-developed, and many results from the classical setting have a nonlinear counterpart. Significant challenges remain in multiple dimensions, and some of what has already been discovered is quite nonintuitive. By answering several classically-inspired questions concerning independence, covariance uncertainty, and behavior under certain linear operations, we continue to highlight the fascinating range of unexpected attributes of the multi-dimensional G-normal distribution.

The material in this chapter is based upon [47], which was presented during the *Financial Mathematics: Advanced Modeling and Numerical Methods* conference at the Université Paris Diderot on June 20, 2014.

1.2 Chapter III Summary

For $\alpha \in (1, 2)$, we present a generalized central limit theorem for α-stable random variables under sublinear expectation. The foundation of our proof is an interior regularity estimate for partial integro-differential equations (PIDEs). A classical generalized central limit theorem is recovered as a special case, provided a mild but natural additional condition holds. Our approach contrasts with previous arguments for the result in the linear setting which have typically relied upon tools that are nonexistent in the sublinear framework, e.g., characteristic functions.

The material in this chapter is based upon [48], which was presented during the following events: the *Financial/Actuarial Mathematics Seminar* at the University of Michigan on December 3, 2014; the *Methods of Mathematical Finance* conference at Carnegie Mellon University on June 2, 2015; the *Mathematical Finance and Probability Seminar* at Rutgers University on November 17, 2015; and at the *Joint Mathematics Meetings: Special Session on Financial Mathematics I* at the Washington State Convention Center on January 7, 2016.

1.3 Chapter IV Summary

We develop a continuous-time game to study the problem faced by an index tracker whose benchmark is undergoing a reconstitution. Given a linear price impact model, we use standard optimality conditions based on the maximum principle to produce candidate Nash equilibria. We analyze these results numerically under varying assumptions regarding tracking error constraints, market characteristics, and predatory trading activity.

The material in this chapter was presented during the *Financial/Actuarial Mathematics Seminar* at the University of Michigan on March 18, 2015.

1.4 Chapter V Summary

How do large-scale participants in parimutuel wagering events affect the house and ordinary bettors? A standard narrative suggests that they may temporarily benefit the former at the expense of the latter. To approach this problem, we begin by developing a model based on the theory of large generalized games. Constrained only by their budgets, a continuum of diffuse (ordinary) players and a single atomic (large-scale) player simultaneously wager to maximize their expected profits according to their individual beliefs. Our main theoretical result gives necessary and sufficient conditions for the existence and uniqueness of a pure-strategy Nash equilibrium. Using this framework, we analyze our question in concrete scenarios. First, we study a situation in which both predicted effects are observed. Neither is always observed in our remaining examples, suggesting the need for a more nuanced view of large-scale participants.

The material in this chapter is based upon [50], which was presented during the *SIAM Conference on Financial Mathematics & Engineering* at the Sheraton Austin Hotel at the Capitol on November 18, 2016. A particular application of these results was also featured in media outlets including the *Associated Press*, after the layman's version [49] was published by the *The Conversation* on May 19, 2016.

1.5 Chapter VI Summary

Oft-cited causes of mini-flash crashes include human errors, endogenous feedback loops, the nature of modern liquidity provision, fundamental value shocks, and market fragmentation. We develop a mathematical model which captures aspects of the first three explanations. Empirical features of recent mini-flash crashes are present in our framework. For example, there are periods when no such events will occur. If they do, even just before their onset, market participants may not know with certainty that a disruption will unfold. Our mini-flash crashes can materialize in both low and high trading volume environments and may be accompanied by a partial synchronization in order submission.

Instead of adopting a classically-inspired equilibrium approach, we borrow ideas from the optimal execution literature. Each of our agents begins with beliefs about how his own trades impact prices and how prices would move in his absence. They, along with other market participants, then submit orders which are executed at a common venue. Naturally, this leads us to explicitly distinguish between how prices actually evolve and our agents' opinions. In particular, every agent's beliefs will be expressly incorrect.

As far as we know, this setup suggests both a new paradigm for modeling heterogeneous agent systems and a novel blueprint for understanding model misspecification risks in the context of optimal execution.

The material in this chapter is based upon [46], which was presented during the *Financial/Actuarial Mathematics Seminar* at the University of Michigan on April 12, 2017.

CHAPTER II

Comparing the G-Normal Distribution to its Classical Counterpart

2.1 Introduction

The G-framework, which includes the G-normal distribution and G-Brownian motion, was initially motivated by the study of risk measures and pricing under volatility uncertainty. Roughly speaking, one can think of these objects as the appropriate analogues of their classical namesakes in a setting of model uncertainty where the relevant collection of probability measures may be singular.

Activity in this area has been considerable since its introduction by Peng ([189], [192]), and developments have proceeded at a rapid pace. A great variety of standard theorems from classical probability and stochastic analysis now have versions in the G-setting including the law of large numbers ([190], [191]), the central limit theorem ([190], [191], [159], [130], [244]), the martingale representation theorem ([223], [224], [195]), Lévy's martingale characterization theorem ([242], [243], [160], [161], [225]), and Girsanov's theorem ([241], [181], [127]). Substantial progress and extensions of this work have been completed in many other directions as well ([176], [93], [135], [177], [94]). Readers interested in survey articles are referred to [188], [193], and [194].

Fundamental issues linger, especially in multiple dimensions. Much of G-stochastic analysis is built upon the G-normal distribution, and yet, many important elementary questions about this distribution remain unanswered. Some of what is known is also rather startling. For instance, the following result about the classical normal distribution is false in the G-setting ([194]):

> For any n-dimensional random vector Z, if $\langle v, Z\rangle$ is a normal random variable for all $v \in \mathbb{R}^n$, then Z is a normal random vector.

Our intuition cannot be trusted when turned to G-normal random vectors. Properties such as the one just mentioned fail due to the nonlinearity of the expectation operator in this framework. Other obstacles include the lack of well-understood tools from the classical theory (e.g., characteristic functions and density functions). Also, the distributional uncertainty associated to a G-normal random vector is far more complex than its initial appearance suggests, since viewing a G-normal random vector as having some fixed but unknown covariance matrix is usually incorrect.

Faced with these challenges, we asked to what extent additional properties of the classical normal distribution hold for its G-counterpart, particularly focusing on behaviors under various linear operations and the relationship between coordinate independence and the covariance matrix. We present our findings concerning the following classical theorems:

(i) Let $Z_1, \dots, Z_n$ be i.i.d. normal random variables. If

$$U = \sum_{i=1}^{n} a_i Z_i \quad \text{and} \quad V = \sum_{i=1}^{n} b_i Z_i$$

for real numbers a_i, b_i satisfying

$$\sum_{i=1}^{n} a_i b_i = 0,$$

then U and V are independent.

(ii) Let $Z_1, \ldots, Z_n$ be independent normal random variables. For any $m \times n$ real matrix A, if

$$Z = (Z_1, \ldots, Z_n)^\top,$$

then AZ is an m-dimensional normal random vector.

(iii) The covariance matrix of a normal random vector is diagonal if and only if its coordinates are (mutually) independent normal random variables.

(iv) If Z is an n-dimensional normal random vector, then there exists an invertible $n \times n$ matrix A such that the coordinates of AZ are independent.

Theorem II.14 reveals that (i) is no longer true in the G-setting. We show in Theorem II.15 that (ii) no longer holds either. While this was already known in a few special cases ([194]), our work explores a broad new class of examples and illuminates a surprising dichotomy depending on the rank of the matrix. Theorems II.17 and II.18 indicate that while (iii) is partially true for the G-normal distribution, unexpected new constraints on the coordinates are introduced. We end by demonstrating with Theorem II.19 that the analogue of (iv) is false.

Our proofs often take advantage of a strange phenomenon in this setting: independence can be asymmetric, i.e., Y can be independent from X even if X is not independent from Y. The general strategy is to show that this is incompatible with the symmetry relations imposed by the G-heat equation associated to the G-normal distribution, given a careful choice of parameters.

These insights expand our knowledge of the remarkable series of behaviors exhibited by the multidimensional G-normal distribution. While the ultimate goal is to use these results to broaden our knowledge of G-stochastic analysis and its related financial applications, we believe that many more surprises lurk in the answers to further theoretical questions about this object.

Readers unfamiliar with this area can find a short treatment of relevant background material in Section 2.2. The specific setup necessary for the statement of our results is in Section 2.3. Our main results are contained in Section 2.4.

2.2 Background

We begin with a brief survey of the theory of sublinear expectation spaces and the G-normal distribution. Our focus will be restricted to only those results that are directly needed for our work in the sequel. Readers interested in a more thorough treatment can find further details in [188], [193], [194], [128], or [163], the references from which our discussion is adapted.

Throughout, we let Ω be a given set and $\mathcal{H}$ be a space of real-valued functions defined on Ω. One should understand $\mathcal{H}$ as a space of random variables on Ω. We will only place minimal emphasis on Ω and $\mathcal{H}$, but we suppose that $\mathcal{H}$

(i) is a linear space,

(ii) contains all constant functions, and

(iii) contains $\varphi(X_1, X_2, \ldots, X_n)$ for every $X_1, X_2, \ldots, X_n \in \mathcal{H}$ and $\varphi \in C_{l.Lip}(\mathbb{R}^n)$, where $C_{l.Lip}(\mathbb{R}^n)$ is the set of functions such that there exists $C > 0$ and $m \in \mathbb{N}$ (depending on φ) satisfying

$$|\varphi(x) - \varphi(y)| \leq C(1 + |x|^m + |y|^m)|x - y|$$

for all x, $y \in \mathbb{R}^n$.

Our specific choice of test functions, $C_{l.Lip}(\mathbb{R}^n)$, is only a matter of convenience. Other spaces are also commonly used.

Definition II.1. A *sublinear expectation* is a function $\hat{\mathbb{E}} : \mathcal{H} \longrightarrow \mathbb{R}$ which is

(i) Monotonic: $\hat{\mathbb{E}}[X] \leq \hat{\mathbb{E}}[Y]$ if $X \leq Y$,

(ii) Constant-preserving: $\hat{\mathbb{E}}[c] = c$ for any $c \in \mathbb{R}$,

(iii) Sub-additive: $\hat{\mathbb{E}}[X+Y] \leq \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y]$, and

(iv) Positive homogeneous: $\hat{\mathbb{E}}[\lambda X] = \lambda \hat{\mathbb{E}}[X]$ for $\lambda \geq 0$.

The triple $\left(\Omega, \mathcal{H}, \hat{\mathbb{E}}\right)$ is called a *sublinear expectation space.*

Notation II.2. Unless stated otherwise, we will always work on some sublinear expectation space $\left(\Omega, \mathcal{H}, \hat{\mathbb{E}}\right)$, and when we have random variables $X_1, X_2, \ldots, X_n \in \mathcal{H}$, we will say that X is an n-dimensional random vector and write $X \in \mathcal{H}^n$.

Great caution is required when manipulating expressions with sublinear expectations due to (iii) and (iv). Most familar operations from the classical theory are simply no longer valid. One situation where a standard technique can be applied is the following.

Lemma II.3. *Consider two random variables $X, Y \in \mathcal{H}$ such that $\hat{\mathbb{E}}[Y] = -\hat{\mathbb{E}}[-Y]$. Then*

$$\hat{\mathbb{E}}[X + \alpha Y] = \hat{\mathbb{E}}[X] + \alpha \hat{\mathbb{E}}[Y]$$

for all $\alpha \in \mathbb{R}$.

In the literature, random variables such as Y above are said to have *no mean-uncertainty*, a notion which also arises in the context of *symmetric G-martingales.* We will resort to a notable consequence of this result again and again: if $\hat{\mathbb{E}}[Y] = \hat{\mathbb{E}}[-Y] = 0$, then for all $\alpha \in \mathbb{R}$,

$$\hat{\mathbb{E}}[X + \alpha Y] = \hat{\mathbb{E}}[X]. \tag{2.1}$$

Definition II.4. An n-dimensional random vector $Y \in \mathcal{H}^n$ is said to be *independent* from an m-dimensional random vector $X \in \mathcal{H}^m$ if for all $\varphi \in C_{l.Lip}\left(\mathbb{R}^{m+n}\right)$, we have

$$\hat{\mathbb{E}}\left[\varphi\left(X, Y\right)\right] = \hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[\varphi\left(x, Y\right)\right]_{x=X}\right].$$

As we mentioned previously, independence can be asymmetric when $\hat{\mathbb{E}}$ is not a linear expectation. A now standard example which we will refer to later illustrates this concretely.

Example II.5. Consider random variables $X, Y \in \mathcal{H}$ such that

(i) $\hat{\mathbb{E}}\left[X\right] = -\hat{\mathbb{E}}\left[-X\right] = 0$,

(ii) $\hat{\mathbb{E}}\left[|X|\right] > 0$, and

(iii) $\hat{\mathbb{E}}\left[Y^2\right] > -\hat{\mathbb{E}}\left[-Y^2\right]$.

If X is independent from Y,

$$\hat{\mathbb{E}}\left[XY^2\right] = 0,$$

while if Y is independent from X,

$$\hat{\mathbb{E}}\left[XY^2\right] > 0.$$

A broad class of situations where (i) , (ii), and (iii) are satisfied occurs when $X \sim \mathcal{N}\left(0, [\underline{\sigma}_1^2, \overline{\sigma}_1^2]\right)$ for $0 < \overline{\sigma}_1^2$ and $Y \sim \mathcal{N}\left(0, [\underline{\sigma}_2^2, \overline{\sigma}_2^2]\right)$ for $\underline{\sigma}_2^2 < \overline{\sigma}_2^2$ (see below for notation).

Ignoring trivial cases, one can actually characterize the distribution of X and Y if X is independent from Y and vice versa (see Proposition II.13 below). Still, observe that if $\hat{\mathbb{E}}$ is a linear expectation, this definition is equivalent to the classical one.

Definition II.6. Let X be an n-dimensional random vector, i.e., $X \in \mathcal{H}^n$.

(i) The *distribution* of X, $\mathbb{F}_X$, is defined on $C_{l.Lip}\left(\mathbb{R}^n\right)$ by

$$\mathbb{F}_X\left(\varphi\right) = \hat{\mathbb{E}}\left[\varphi\left(X\right)\right].$$

(ii) X has *distributional uncertainty* if $\mathbb{F}_X$ is not a linear expectation ($(\mathbb{R}^n, C_{l.lip}(\mathbb{R}^n), \mathbb{F}_X)$ is always a sublinear expectation space).

(iii) If $Y \in \mathcal{H}^n$ is another n-dimensional random vector, then X and Y are *identically distributed*, denoted $X \sim Y$, if

$$\mathbb{F}_X = \mathbb{F}_Y.$$

(iv) If X and Y are identically distributed and Y is independent from X, then Y is an *independent copy* of X.

This notion of being identically distributed is also equivalent to the classical defintion if $\hat{\mathbb{E}}$ is a linear expectation. The sublinear case possesses many interesting new features, but we will only need to know about those pertaining to the *G-normal distribution* and the *maximal distribution*.

Definition II.7. An n-dimensional random vector $X \in \mathcal{H}^n$ is said to be *G-normally distributed* if for any independent copy of X, say $\bar{X}$, we have

$$aX + b\bar{X} \sim \sqrt{a^2 + b^2}X$$

for all $a, b \geq 0$.

"G" refers to the sublinear function defined on the space of $n \times n$ symmetric matrices, $\mathbb{S}(n)$, by

$$G(A) = \frac{1}{2}\hat{\mathbb{E}}[\langle AX, X\rangle]. \tag{2.2}$$

For each such function, there exists a unique bounded, closed, convex subset Γ of the $n \times n$ positive semidefinite matrices, $\mathbb{S}^+(n)$, such that

$$G(A) = \frac{1}{2}\sup_{B\in\Gamma} \operatorname{tr}[AB]. \tag{2.3}$$

Conversely, given any Γ with these properties, there exists a G-normal random vector X such that (2.2) and (2.3) hold.

Γ completely determines the distribution of a G-normal random vector X, and in fact, one can loosely interpret Γ as describing the covariance uncertainty of X. As we remarked above, some care must be exercised since viewing X as possessing a classical normal distribution with some fixed but unknown covariance matrix selected from Γ is not generally correct. However, if Γ contains only one element, then X is a classical normal random vector with mean zero and covariance Γ.

Notation II.8. We write $X \sim \mathcal{N}(0, \Gamma)$. If $n = 1$, then $\Gamma = [\underline{\sigma}^2, \overline{\sigma}^2]$ for $0 \leq \underline{\sigma}^2 \leq \overline{\sigma}^2$ given by

$$\overline{\sigma}^2 = \hat{\mathbb{E}}\left[X^2\right] \quad \text{and} \quad \underline{\sigma}^2 = -\hat{\mathbb{E}}\left[-X^2\right].$$

We will frequently use the following important basic properties of a G-normal random vector.

Lemma II.9. *Let $X = (X_1, \ldots, X_n)$ be an n-dimensional G-normal random vector, i.e., $X \sim \mathcal{N}(0, \Gamma)$. Then*

(i) $\hat{\mathbb{E}}[X_i] = \hat{\mathbb{E}}[-X_i] = 0$ for all i;

(ii) $-X$ and X are identically distributed, i.e., $-X \sim \mathcal{N}(0, \Gamma)$;

(iii) for all $v \in \mathbb{R}^n$, the random variable $\langle v, X\rangle$ is $\mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$-distributed, where

$$\overline{\sigma}^2 = \hat{\mathbb{E}}\left[\langle v, X\rangle^2\right] \quad \textit{and} \quad \underline{\sigma}^2 = -\hat{\mathbb{E}}\left[-\langle v, X\rangle^2\right] \textit{; and}$$

(iv) for all $m \times n$ real matrices M, $MX \sim \mathcal{N}\left(0, M\Gamma M^\top\right)$.

Perhaps the deepest known property of a G-normal random vector is its intimate connection to the so-called "G-heat equation", a parabolic PDE generalizing the classical heat equation.

Proposition II.10. *Let X be an n-dimensional G-normal random vector, i.e., $X \sim \mathcal{N}(0,\Gamma)$. For all $\varphi \in C_{l.Lip}(\mathbb{R}^n)$, the function*

$$u(t,x) = \hat{\mathbb{E}}\left[\varphi\left(x+\sqrt{t}X\right)\right], \quad (t,x) \in [0,\infty) \times \mathbb{R}^n$$

is the unique viscosity solution of the following PDE:

$$\partial_t u - G\left(D^2 u\right) = 0, \quad (t,x) \in (0,\infty) \times \mathbb{R}^n$$
$$u(0,x) = \varphi(x), \quad x \in \mathbb{R}^n,$$

where $D^2 u = \left(\partial^2_{x^i x^j} u\right)_{ij}$.

We will not need the full theory of viscosity solutions in our development, as the solution to the equation above is actually a classical solution if G is non-degenerate.

We will use the remainder of this section only for the proof of Theorem II.18.

Notation II.11. We let $C_{b.Lip}(\mathbb{R}^n)$ denote the space of bounded Lipschitz functions on $\mathbb{R}^n$.

We recall that [128] uses $C_{b.Lip}(\mathbb{R}^n)$ as the space of test functions instead of $C_{l.Lip}(\mathbb{R}^n)$. This technicality does not matter for our proof of Theorem II.18, as we will explain later.

Definition II.12. An n-dimensional random vector $X \in \mathcal{H}^n$ is called *maximally distributed* if there exists a closed set $\Gamma \subset \mathbb{R}^n$ such that

$$\hat{\mathbb{E}}[\varphi(X)] = \sup_{x\in\Gamma} \varphi(x)$$

for all $\varphi \in C_{b.Lip}(\mathbb{R}^n)$.

One can understand random variables of this kind as analogues of constants in the sublinear setting. In particular, if X is a G-normal random variable with $X \sim$

$\mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$ for some $\underline{\sigma}^2 < \overline{\sigma}^2$, then X is not maximally distributed. Observe that we have presented the definition from [128] rather than the original version in [194], as we will need the following theorem from [128].

Proposition II.13. *Suppose that the random variable $W \in \mathcal{H}$ has distributional uncertainty and that the random variable $W' \in \mathcal{H}$ is not a constant. If W is independent from W' and vice versa, then W and W' must be maximally distributed.*

2.3 Basic Setup

Throughout our work below, we consider random vectors $X, Y \in \mathcal{H}^n$ such that

(i) $X = (X_1, \ldots, X_n)$ is a G-normal random vector, i.e., $X \sim \mathcal{N}(0, \Gamma)$; and

(ii) $Y = (Y_1, \ldots, Y_n)$, where

  (a) $Y_1 \sim \mathcal{N}(0, [\underline{\sigma}^2, \overline{\sigma}^2])$ for some $0 < \underline{\sigma}^2 < \overline{\sigma}^2$;

  (b) Y_{i+1} and Y_i are identically distributed, i.e., $Y_{i+1} \sim Y_i$; and

  (c) Y_{i+1} is independent from $(Y_1, \ldots, Y_i)$.

Although a seemingly innocent construction, we will see that Y is a rich source of counterexamples when evaluating whether or not standard classical theorems about the normal distribution hold in the G-framework. The inequality

$$\underline{\sigma}^2 < \overline{\sigma}^2$$

is critical for us, as it implies that the coordinates of Y are not classical normal random variables. Because our objective is to compare the properties of the classical normal distribution with its G-counterpart, we have added the assumption

$$0 < \underline{\sigma}^2$$

for convenience. As observed above, this condition ensures that the solution to the G-heat equation is actually classical.

Whenever restating classical results to facilitate our comparisons, we will always call classical random vectors "Z".

2.4 Main Results

2.4.1 Behavior Under Linear Combinations

Recall that in the classical setting, we have the following result:

Let $Z_1, \ldots, Z_n$ be i.i.d. normal random variables. If

$$U = \sum_{i=1}^{n} a_i Z_i \quad \text{and} \quad V = \sum_{i=1}^{n} b_i Z_i$$

for real numbers a_i, b_i satisfying

$$\sum_{i=1}^{n} a_i b_i = 0,$$

then U and V are independent.

The corresponding statement does not hold in the G-framework.

Theorem II.14. *Let $a_1 = 1$, $a_2 = 1$, $b_1 = 1$, and $b_2 = -1$. Set the remaining constants equal to zero, i.e.,*

$$U = Y_1 + Y_2 \quad \textit{and} \quad V = Y_1 - Y_2.$$

Then

$$\sum_{i=1}^{n} a_i b_i = 0,$$

but U is not independent from V and vice versa.

In fact, it is not yet known if any non-trivial linear combination of this kind will produce independent random variables. An important classical characterization of the normal distribution, the Skitovich-Darmois theorem, is related to the independence of such linear combinations, and whether or not a G-version of this holds is also unknown.

Proof. Our strategy will be to show that the random vectors (U, V) and (V, U) are identically distributed. On the other hand, if U were independent from V or vice versa, the resulting destruction of symmetry would make this impossible.

A simple calculation shows that $-Y_2$ is an independent copy of Y_1, which means that

$$V \sim \sqrt{2} Y_1 \sim \mathcal{N}\left(0, \left[2\underline{\sigma}^2, 2\overline{\sigma}^2\right]\right)$$

by Definition II.7. The same is true for U.

By Example II.5, if U is independent from V, then

$$\hat{\mathbb{E}}\left[UV^2\right] = 0$$

and

$$\hat{\mathbb{E}}\left[VU^2\right] > 0.$$

If V is independent from U, then

$$\hat{\mathbb{E}}\left[VU^2\right] = 0$$

and

$$\hat{\mathbb{E}}\left[UV^2\right] > 0.$$

Hence, exactly one of these must hold:

(i) U is independent from V but V is not independent from U.

(ii) V is independent from U but U is not independent from V.

(iii) U is not independent from V and V is not independent from U.

Let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$S(x, y) = (x - y, x + y) \quad \text{for all } (x, y) \in \mathbb{R}^2.$$

Observe that $\varphi \circ S \in C_{l.Lip}(\mathbb{R}^2)$ for any $\varphi \in C_{l.Lip}(\mathbb{R}^2)$. Since Y_2 and $-Y_2$ are each independent copies of Y_1, for any $\varphi \in C_{l.Lip}(\mathbb{R}^2)$,

$$\begin{aligned}
\hat{\mathbb{E}}[\varphi(V, U)] &= \hat{\mathbb{E}}[\varphi(Y_1 - Y_2, Y_1 + Y_2)] \\
&= \hat{\mathbb{E}}[(\varphi \circ S)(Y_1, Y_2)] \\
&= \hat{\mathbb{E}}\left[\hat{\mathbb{E}}[(\varphi \circ S)(\bar{x}, Y_2)]_{\bar{x}=Y_1}\right] \\
&= \hat{\mathbb{E}}\left[\hat{\mathbb{E}}[(\varphi \circ S)(\bar{x}, -Y_2)]_{\bar{x}=Y_1}\right] \\
&= \hat{\mathbb{E}}[(\varphi \circ S)(Y_1, -Y_2)] \\
&= \hat{\mathbb{E}}[\varphi(Y_1 + Y_2, Y_1 - Y_2)] \\
&= \hat{\mathbb{E}}[\varphi(U, V)].
\end{aligned}$$

Applying this equality to $\varphi(x, y) = xy^2$, we have

$$\hat{\mathbb{E}}\left[VU^2\right] = \hat{\mathbb{E}}\left[UV^2\right],$$

which implies that U is not independent from V and vice versa. $\square$

2.4.2 Behavior Under General Linear Transformations

Allowing the degenerate case, we have another important property of the classical normal distribution:

> Let $Z_1, \ldots, Z_n$ be independent normal random variables. For any $m \times n$ real matrix A, if
>
> $$Z = (Z_1, \ldots, Z_n)^\top,$$
>
> then AZ is an m-dimensional normal random vector.

The situation in the G-framework is far more delicate.

Theorem II.15. *For any $m \times n$ real matrix A,*

(i) $\langle v, AY \rangle$ is a G-normal random variable with

$$\langle v, AY \rangle \sim \mathcal{N}\left(0, \left[\left\|v^\top A\right\|^2 \underline{\sigma}^2, \left\|v^\top A\right\|^2 \overline{\sigma}^2\right]\right)$$

for all $v \in \mathbb{R}^m$;

(ii) if A has rank less than or equal to one, AY is an m-dimensional G-normal random vector, or more precisely, $AY \sim \mathcal{N}(0, \Gamma')$, where

$$\Gamma' = \left\{uru^\top : r \in \left[\|w\|^2 \underline{\sigma}^2, \|w\|^2 \overline{\sigma}^2\right]\right\}$$

and $A = uw^\top$ for $u \in \mathbb{R}^m$, $w \in \mathbb{R}^n$; and

(iii) if A is invertible, AY is not G-normally distributed. (In particular, Y is not G-normally distributed.)

One would expect (i) and (ii); however, (iii) is quite surprising both because of the bifurcation it reveals and its relation to classical theorems. While it was previously known that the classical property above failed to be true in the G-setting ([194]), our result provides an expansive new series of cases illustrating this failure. It serves the same purpose with respect to the following classical statement as well:

> For any n-dimensional random vector Z, if $\langle v, Z \rangle$ is a normal random variable for all $v \in \mathbb{R}^n$, then Z is an n-dimensional normal random vector.

Whether or not AY is a G-normal random vector if A is non-invertible but has rank strictly greater than one remains unclear.

The most difficult part of the proof is the following lemma, which is a small extension of Exercise 1.15 in [194]. This lemma will be critical for our proof of Theorems II.17, II.18, and II.19 as well.

Lemma II.16. *Let $\alpha > 0$. Suppose that W_1 and W_2 are two G-normal random variables such that*

(i) $W_1 \sim \mathcal{N}\left(0, [\underline{\sigma}^2, \overline{\sigma}^2]\right)$*; and*

(ii) $W_2 \sim \mathcal{N}\left(0, [\alpha\underline{\sigma}^2, \alpha\overline{\sigma}^2]\right)$.

If either W_2 is independent from W_1 or vice versa, then $W = (W_1, W_2)^\top$ is not a 2-dimensional G-normal random vector.

Proof. We first consider the case where W_2 is independent from W_1. Suppose instead that W is a 2-dimensional G-normal random vector. Our initial step will be to compute the corresponding G-heat equation, which we will use to establish an identity relating the distributions of the random vectors $W = (W_1, W_2)^\top$ and $(W_2, W_1)^\top$. The conclusion will be reached by showing that this "symmetry" contradicts the asymmetry induced by our independence assumption.

We can find a bounded, closed, convex subset $\Gamma \subset \mathbb{S}^+(2)$ such that

$$\frac{1}{2}\hat{\mathbb{E}}\left[\langle AW, W\rangle\right] = G(A) = \frac{1}{2}\sup_{B\in\Gamma} \operatorname{tr}[AB]$$

for all $A \in \mathbb{S}(2)$. Now for all

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \in \mathbb{S}(2),$$

we have

$$
\begin{aligned}
G(A) &= \frac{1}{2}\hat{\mathbb{E}}\left[\left\langle \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}, \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \right\rangle\right] \\
&= \frac{1}{2}\hat{\mathbb{E}}\left[a_{11}W_1^2 + 2a_{12}W_1W_2 + a_{22}W_2^2\right] \\
&= \frac{1}{2}\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[a_{11}\bar{x}^2 + 2a_{12}\bar{x}W_2 + a_{22}W_2^2\right]_{\bar{x}=W_1}\right] \\
&= \frac{1}{2}\hat{\mathbb{E}}\left[\left(a_{11}\bar{x}^2 + 2a_{12}\bar{x}\hat{\mathbb{E}}\left[W_2\right] + \hat{\mathbb{E}}\left[a_{22}W_2^2\right]\right)_{\bar{x}=W_1}\right] \\
&= \frac{1}{2}\hat{\mathbb{E}}\left[\left(a_{11}\bar{x}^2 + \alpha\overline{\sigma}^2\left(a_{22}^+\right) - \alpha\underline{\sigma}^2\left(a_{22}^-\right)\right)_{\bar{x}=W_1}\right] \\
&= \frac{1}{2}\hat{\mathbb{E}}\left[a_{11}W_1^2\right] + \frac{\alpha}{2}\left(\overline{\sigma}^2\left(a_{22}^+\right) - \underline{\sigma}^2\left(a_{22}^-\right)\right) \\
&= \frac{1}{2}\left(\overline{\sigma}^2\left(a_{11}^+\right) - \underline{\sigma}^2\left(a_{11}^-\right)\right) + \frac{\alpha}{2}\left(\overline{\sigma}^2\left(a_{22}^+\right) - \underline{\sigma}^2\left(a_{22}^-\right)\right) \\
&= \bar{G}\left(a_{11}\right) + \alpha\bar{G}\left(a_{22}\right),
\end{aligned}
$$

where $\bar{G}$ is defined by

$$\bar{G}(x) = \frac{1}{2}\left(\overline{\sigma}^2\left(x^+\right) - \underline{\sigma}^2\left(x^-\right)\right)$$

for all $x \in \mathbb{R}$. A quick calculation verifies that Γ must then be given by

$$\Gamma = \left\{ \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} : r_1 \in \left[\underline{\sigma}^2, \overline{\sigma}^2\right], r_2 \in \left[\alpha\underline{\sigma}^2, \alpha\overline{\sigma}^2\right] \right\}.$$

Let $\varphi \in C_{l.Lip}\left(\mathbb{R}^2\right)$. Define the function u by

$$u(t,x,y) = \hat{\mathbb{E}}\left[\varphi\left((x,y) + \sqrt{t}\left(W_1, W_2\right)\right)\right]$$

for all $(t,x,y) \in [0,\infty) \times \mathbb{R}^2$. By Proposition II.10, u is the unique viscosity solution to

$$
\begin{aligned}
\partial_t u - \bar{G}\left(\partial_{xx}^2 u\right) - \alpha\bar{G}\left(\partial_{yy}^2 u\right) = 0, &\quad (t,x,y) \in (0,\infty) \times \mathbb{R}^2 \\
u(0,x,y) = \varphi(x,y), &\quad (x,y) \in \mathbb{R}^2.
\end{aligned}
$$

In fact, since $0 < \underline{\sigma}^2$, u is a classical solution.

Let the functions S and $\tilde{S}$ be given by

$$S(x,y) = \left(\frac{1}{\sqrt{\alpha}}y, \sqrt{\alpha}x\right)$$

for all $(x,y) \in \mathbb{R}^2$ and

$$\tilde{S}(t,x,y) = \left(t, \frac{1}{\sqrt{\alpha}}y, \sqrt{\alpha}x\right)$$

for all $(t,x,y) \in [0,\infty) \times \mathbb{R}^2$. Define a function v by

$$v = u \circ \tilde{S}.$$

Then v is a classical solution of

$$\begin{aligned} \partial_t v - \bar{G}\left(\partial_{xx}^2 v\right) - \alpha\bar{G}\left(\partial_{yy}^2 v\right) = 0, \quad (t,x,y) \in (0,\infty) \times \mathbb{R}^2 \\ v(0,x,y) = (\varphi \circ S)(x,y), \quad (x,y) \in \mathbb{R}^2. \end{aligned}$$

Hence,

$$v(t,x,y) = \hat{\mathbb{E}}\left[(\varphi \circ S)\left((x,y) + \sqrt{t}(W_1, W_2)\right)\right].$$

In particular,

$$\begin{aligned} &\hat{\mathbb{E}}\left[\varphi\left(\frac{1}{\sqrt{\alpha}}W_2, \sqrt{\alpha}W_1\right)\right] \\ &= \hat{\mathbb{E}}\left[(\varphi \circ S)\left((0,0) + \sqrt{1}(W_1, W_2)\right)\right] \\ &= v(1,0,0) \\ &= \left(u \circ \tilde{S}\right)(1,0,0) \\ &= u(1,0,0) \\ &= \hat{\mathbb{E}}\left[\varphi\left((0,0) + \sqrt{1}(W_1, W_2)\right)\right] \\ &= \hat{\mathbb{E}}\left[\varphi(W_1, W_2)\right]. \end{aligned}$$

Since $\varphi \in C_{l.Lip}(\mathbb{R}^2)$ was arbitrary, applying this to the function $\varphi(x,y) = xy^2$ gives

$$\sqrt{\alpha}\hat{\mathbb{E}}\left[W_2 W_1^2\right] = \hat{\mathbb{E}}\left[W_1 W_2^2\right].$$

On the other hand, Example II.5 implies

$$\hat{\mathbb{E}}\left[W_2 W_1^2\right] = 0$$

and

$$\hat{\mathbb{E}}\left[W_1 W_2^2\right] > 0,$$

a contradiction. It follows that W is not a 2-dimensional G-normal random vector in this case.

To finish the proof, assume that W_1 is independent from W_2. If $W = (W_1, W_2)^\top$ were 2-dimensional G-normal random vector, then

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} W_2 \\ W_1 \end{bmatrix}$$

would be a 2-dimensional G-normal random vector by Lemma II.9. This is impossible by what we just considered, so the result holds. □

The proof of the theorem is now straightforward.

Proof. Let $A = (a_{ij})$. To prove (i), suppose $v = (v_1, \dots, v_m)^\top \in \mathbb{R}^m$. For all $\varphi \in C_{l.Lip}(\mathbb{R})$,

$$\begin{aligned}
\hat{\mathbb{E}}\left[\varphi\left(\langle v, AY\rangle\right)\right] &= \hat{\mathbb{E}}\left[\varphi\left(\left(v_1 \sum_{k=1}^{n} a_{1k} Y_k\right) + \cdots + \left(v_m \sum_{k=1}^{n} a_{mk} Y_k\right)\right)\right] \\
&= \hat{\mathbb{E}}\left[\varphi\left(\left(\sum_{k=1}^{m} v_k a_{k1}\right) Y_1 + \cdots + \left(\sum_{k=1}^{m} v_k a_{kn}\right) Y_n\right)\right] \\
&= \hat{\mathbb{E}}\left[\varphi\left(\sqrt{\left(\sum_{k=1}^{m} v_k a_{k1}\right)^2 + \cdots + \left(\sum_{k=1}^{m} v_k a_{kn}\right)^2} Y_1\right)\right] \\
&= \hat{\mathbb{E}}\left[\varphi\left(\left\|v^\top A\right\| Y_1\right)\right],
\end{aligned}$$

which directly follows from Definition II.7. By Lemma II.9,

$$\langle v, AY\rangle \sim \left\|v^\top A\right\| Y_1 \sim \mathcal{N}\left(0, \left[\left\|v^\top A\right\|^2 \underline{\sigma}^2, \left\|v^\top A\right\|^2 \overline{\sigma}^2\right]\right).$$

For (ii), since A has rank less than or equal to one, we can find $u = (u_1, \ldots, u_m)^\top \in \mathbb{R}^m$ and $w = (w_1, \ldots, w_n)^\top \in \mathbb{R}^n$ such that

$$A = uw^\top.$$

By Lemma II.9,

$$w^\top Y \sim \left(\sqrt{\sum_{i=1}^n w_i^2}\right) Y_1 \sim \mathcal{N}\left(0, \left[\|w\|^2 \underline{\sigma}^2, \|w\|^2 \overline{\sigma}^2\right]\right).$$

Since AY is given by

$$AY = u\left(w^\top Y\right),$$

another application of Lemma II.9 implies the result.

To see that (iii) holds, let B be the $2 \times n$ matrix

$$B = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{bmatrix}.$$

Since

$$\left(BA^{-1}\right) AY = BY = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix},$$

Lemmas II.9 and II.16 show that AY cannot be an n-dimensional G-normal random vector. □

2.4.3 Connections Between Covariance Uncertainty and Independence

Classically, there is a tight relationship between the covariance matrix of a normal random vector and the independence of its coordinates:

> The covariance matrix of a normal random vector is diagonal if and only if its coordinates are (mutually) independent normal random variables.

Once again, the analogous situation is more subtle in the G-setting. For instance, the forward direction is false.

Theorem II.17. *Let*

$$\Gamma = \left\{ \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} : r_1, r_2 \in \left[\underline{\sigma}^2, \overline{\sigma}^2\right] \right\},$$

so $X = (X_1, X_2)^\top$ *and* $X \sim \mathcal{N}(0, \Gamma)$. X_1 *is not independent from* X_2 *and vice versa.*

Proof. We will proceed by computing the distributions of the random variables X_1 and X_2. Then we will invoke Lemma II.16, the impetus for our choice of this specific Γ.

Recall that for $A \in \mathbb{S}(2)$,

$$\frac{1}{2}\hat{\mathbb{E}}\left[\langle AX, X\rangle\right] = G(A) = \frac{1}{2}\sup_{B\in\Gamma} \operatorname{tr}[AB].$$

In particular,

$$\begin{aligned} \frac{1}{2}\hat{\mathbb{E}}\left[X_1^2\right] &= G\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) \\ &= \frac{1}{2}\sup_{B\in\Gamma} \operatorname{tr}\left[\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} B\right] \\ &= \frac{1}{2}\sup_{r_1\in\left[\underline{\sigma}^2,\overline{\sigma}^2\right]} r_1 \\ &= \frac{1}{2}\overline{\sigma}^2, \end{aligned}$$

i.e.,

$$\hat{\mathbb{E}}\left[X_1^2\right] = \overline{\sigma}^2.$$

A similar calculation gives that

$$\hat{\mathbb{E}}\left[X_2^2\right] = \overline{\sigma}^2$$

and

$$-\hat{\mathbb{E}}\left[-X_1^2\right] = -\hat{\mathbb{E}}\left[-X_2^2\right] = \underline{\sigma}^2.$$

By Lemma II.9, X_1 and X_2 are both G-normal random variables and

$$X_1, X_2 \sim \mathcal{N}\left(0, \left[\underline{\sigma}^2, \overline{\sigma}^2\right]\right).$$

Lemma II.16 implies that X_1 cannot be independent from X_2 and vice versa. $\square$

The backward direction bears a stronger resemblance to the classical case, although with a few unforeseen twists.

Theorem II.18. *Suppose that there exists a permutation $\pi \in S_n$ such that for all $1 \leq i \leq n-1$, $X_{\pi(i+1)}$ is independent from $\left(X_{\pi(1)}, \ldots, X_{\pi(i)}\right)$. Then*

(i) for $\underline{\sigma}_i^2$, $\overline{\sigma}_i^2$ such that $X_i \sim \mathcal{N}\left(0, [\underline{\sigma}_i^2, \overline{\sigma}_i^2]\right)$,

$$\Gamma = \left\{ diag\left[r_1, \ldots, r_n\right] : r_i \in \left[\underline{\sigma}_i^2, \overline{\sigma}_i^2\right] \right\};$$

(ii) for any $1 \leq i \leq n$ such that $0 < \underline{\sigma}_i^2 < \overline{\sigma}_i^2$,

$$\alpha\left[\underline{\sigma}_i^2, \overline{\sigma}_i^2\right] \neq \left[\underline{\sigma}_j^2, \overline{\sigma}_j^2\right]$$

for all $j \neq i$ and $\alpha > 0$; and

(iii) for any $i < j$, $X_{\pi(i)}$ is not independent from $X_{\pi(j)}$ if either of the following hold:

(a) $\underline{\sigma}_{\pi(i)}^2 < \overline{\sigma}_{\pi(i)}^2$ and $0 < \overline{\sigma}_{\pi(j)}^2$, or

(b) $\underline{\sigma}_{\pi(j)}^2 < \overline{\sigma}_{\pi(j)}^2$ and $0 < \overline{\sigma}_{\pi(i)}^2$.

(i) is exactly as expected, but (ii) and (iii) are highly nonintuitive: no remotely similar conditions are present in the classical theory. Observe that while this theorem places substantial restrictions on Γ if the coordinates of X satisfy appropriate independence conditions, it does not address the existence of such an X. This issue is still unresolved.

Proof. We begin with (i). By Lemma II.9 we can find $0 \leq \underline{\sigma}_i^2 \leq \overline{\sigma}_i^2$ such that

$$X_i \sim \mathcal{N}\left(0, \left[\underline{\sigma}_i^2, \overline{\sigma}_i^2\right]\right)$$

for all $1 \leq i \leq n$. Also, for any $A = (a_{ij}) \in \mathbb{S}(n)$,

$$\frac{1}{2}\hat{\mathbb{E}}\left[\langle AX, X\rangle\right] = G(A) = \frac{1}{2}\sup_{B\in\Gamma}\operatorname{tr}[AB].$$

Now

$$\begin{aligned}\frac{1}{2}\hat{\mathbb{E}}\left[\langle AX, X\rangle\right] &= \frac{1}{2}\hat{\mathbb{E}}\left[\sum_{i=1}^{n} a_{ii}X_i^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_{ij}X_iX_j\right] \\ &= \frac{1}{2}\hat{\mathbb{E}}\left[\sum_{i=1}^{n} a_{\pi(i)\pi(i)}X_{\pi(i)}^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_{\pi(i)\pi(j)}X_{\pi(i)}X_{\pi(j)}\right].\end{aligned}$$

For any $1 \leq i < j \leq n$,

$$\hat{\mathbb{E}}\left[X_{\pi(i)}X_{\pi(j)}\right] = \hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[\bar{x}X_{\pi(j)}\right]_{\bar{x}=X_{\pi(i)}}\right] = 0$$

and

$$-\hat{\mathbb{E}}\left[-X_{\pi(i)}X_{\pi(j)}\right] = -\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[-\bar{x}X_{\pi(j)}\right]_{\bar{x}=X_{\pi(i)}}\right] = 0.$$

By repeated application of (2.1),

$$
\begin{aligned}
\frac{1}{2}\hat{\mathbb{E}}\left[\langle AX, X\rangle\right] &= \frac{1}{2}\hat{\mathbb{E}}\left[\sum_{i=1}^{n} a_{\pi(i)\pi(i)} X_{\pi(i)}^2\right] \\
&= \frac{1}{2}\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[\cdots\hat{\mathbb{E}}\left[\sum_{i=1}^{n-1} a_{\pi(i)\pi(i)} \bar{x}_{\pi(i)}^2\right.\right.\right. \\
&\qquad \left.\left.\left. + a_{\pi(n)\pi(n)} X_{\pi(n)}^2\right]_{\bar{x}_{\pi(n-1)}=X_{\pi(n-1)}} \cdots\right]_{\bar{x}_{\pi(1)}=X_{\pi(1)}}\right] \\
&= \frac{1}{2}\sum_{i=1}^{n}\left(\overline{\sigma}_{\pi(i)}^2\left(a_{\pi(i)\pi(i)}^+\right) - \underline{\sigma}_{\pi(i)}^2\left(a_{\pi(i)\pi(i)}^-\right)\right) \\
&= \frac{1}{2}\sum_{i=1}^{n}\left(\overline{\sigma}_i^2\left(a_{ii}^+\right) - \underline{\sigma}_i^2\left(a_{ii}^-\right)\right).
\end{aligned}
$$

Hence, we need to find a bounded, closed, convex subset $\Gamma \subset \mathbb{S}^+(n)$ such that

$$\sum_{i=1}^{n}\left(\overline{\sigma}_i^2\left(a_{ii}^+\right) - \underline{\sigma}_i^2\left(a_{ii}^-\right)\right) = \sup_{B\in\Gamma} \operatorname{tr}[AB]$$

for any $A = (a_{ij}) \in \mathbb{S}(n)$. One easily verifies that

$$\Gamma = \left\{\operatorname{diag}[r_1, \ldots, r_n] : r_i \in \left[\underline{\sigma}_i^2, \overline{\sigma}_i^2\right]\right\}.$$

To prove (ii), suppose that $i \neq j$ and $0 < \underline{\sigma}_i^2 < \overline{\sigma}_i^2$. Let B_{ij} be the $2 \times n$ matrix of 0's and 1's such that

$$\begin{bmatrix} X_i \\ X_j \end{bmatrix} = B_{ij}X.$$

Lemma II.9 implies that

$$\begin{bmatrix} X_i \\ X_j \end{bmatrix}$$

is a 2-dimensional G-normal random vector. By Lemma II.16, since either X_j is independent from X_i or vice versa, (ii) holds.

(iii) is an immediate consequence of Proposition II.13. One might object that the space of test functions in [128] is $C_{b.Lip}(\mathbb{R}^n)$ instead of $C_{l.Lip}(\mathbb{R}^n)$, but this issue is addressed in Example 21 of that reference. $\square$

An additional classical result bridging the form of a normal random vector's covariance matrix and the independence of its coordinates is as follows:

> If Z is an n-dimensional normal random vector, then there exists an invertible $n \times n$ matrix A such that the coordinates of AZ are independent.

The related statement is false for a G-normal random vector. There are several possible approaches here. For example, by Lemma II.9 and Theorem II.18, it suffices to construct a Γ such that $A\Gamma A^\top$ contains non-diagonal matrices for all invertible $n \times n$ matrices A. Our method uses a more straightforward choice of Γ but also incorporates Lemma II.16.

Theorem II.19. *Let*

$$\Gamma = \left\{ \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} : r_1, r_2 \in \left[\underline{\sigma}^2, \overline{\sigma}^2\right] \right\}.$$

There is no invertible 2×2 real matrix A such that the coordinates of AX are independent.

Proof. Suppose that for some invertible 2×2 real matrix $A = (a_{ij})$, either W_2 is independent from W_1 or vice versa, where $(W_1, W_2)^\top = AX$. By Lemma II.9,

$$AX \sim \mathcal{N}\left(0, A\Gamma A^\top\right).$$

For all $r_1, r_2 \in [\underline{\sigma}^2, \overline{\sigma}^2]$,

$$A \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} A^\top = \begin{bmatrix} r_1 a_{11}^2 + r_2 a_{12}^2 & r_1 a_{11} a_{21} + r_2 a_{12} a_{22} \\ r_1 a_{11} a_{21} + r_2 a_{12} a_{22} & r_1 a_{21}^2 + r_2 a_{22}^2 \end{bmatrix}.$$

From Theorem II.18,

$$A\Gamma A^\top = \left\{ \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} : r_1 \in \left[\underline{\sigma}_1^2, \overline{\sigma}_1^2\right], r_2 \in \left[\underline{\sigma}_2^2, \overline{\sigma}_2^2\right] \right\}$$

where $W_i \sim \mathcal{N}\left(0, [\underline{\sigma}_i^2, \overline{\sigma}_i^2]\right)$.

This is only possible if

$$a_{11}a_{21} = a_{12}a_{22} = 0.$$

Since A is an invertible 2×2 real matrix, it must be of the form

$$A = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \quad \text{for nonzero } \ a_{11}, a_{22} \in \mathbb{R}$$

or

$$A = \begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix} \quad \text{for nonzero } \ a_{12}, a_{21} \in \mathbb{R}.$$

If the former holds, then

$$A\Gamma A^\top = \left\{ \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} : r_1 \in \left[a_{11}^2\underline{\sigma}^2, a_{11}^2\overline{\sigma}^2\right], r_2 \in \left[a_{22}^2\underline{\sigma}^2, a_{22}^2\overline{\sigma}^2\right] \right\}.$$

In this case,

$$W_1 \sim \mathcal{N}\left(0, \left[a_{11}^2\underline{\sigma}^2, a_{11}^2\overline{\sigma}^2\right]\right) \quad \text{and} \quad W_2 \sim \mathcal{N}\left(0, \left[a_{22}^2\underline{\sigma}^2, a_{22}^2\overline{\sigma}^2\right]\right)$$

by Lemma II.9, which is impossible by Lemma II.16.

Similarly, A cannot have the latter form, so the result holds. $\square$

CHAPTER III

An α-Stable Limit Theorem Under Sublinear Expectation

3.1 Introduction

The purpose of this manuscript is to prove a generalized central limit theorem for α-stable random variables in the setting of sublinear expectation. Such a result complements the limit theorems for G-normal random variables due to Peng and others in this context and answers in the affirmative a question posed by Neufeld and Nutz in [173] (see below).

When working with a sublinear expectation, one is simultaneously considering a potentially uncountably infinite and non-dominated collection of probability measures. A construction of this kind is motivated by the study of pricing under volatility uncertainty. Needless to say, a variety of frequently called upon devices from the classical setting are unavailable. The complications encompass further issues as well: new behaviors are occasionally observed like those outlined in [47].

Analogues of significant theorems from classical probability and stochastic analysis are nevertheless moderately abundant. For instance, versions of the law of large numbers can be found in [190] and [191]; the martingale representation theorem is given in [223], [224], and [195]; Girsanov's theorem is obtained in [241], [181], and [127]; and a Donsker-type result is shown in [94]. To conduct investigations along these lines, standard proofs must often be reimagined. For instance, Peng's proof of the central limit theorem under sublinear expectation in [190] resorts to interior regularity estimates for fully nonlinear parabolic partial differential equations (PDEs). His idea has since been extended to prove a number of variants of his original result, e.g., see [191], [159], [130], and [244].

We will operate in the sublinear expectation framework unless explicitly indicated otherwise. The objects of our special attention here, the α-stable random variables for $\alpha \in (1, 2)$, were introduced in [173]. The authors pondered whether or not these could be the subject of a generalized central limit theorem. Classical generalized central limit theorems ordinarily come in one of three flavors:

(i) a statement indicating that a random variable has a nonempty domain of attraction if and only if it is α-stable such as Theorem 2.1.1 in [136],

(ii) a characterization theorem for the domain of attraction of an α-stable random variable such as Theorem 2.6.1 in [136], or

(iii) a characterization theorem for the domain of normal attraction for an α-stable random variable such as Theorem 2.6.7 in [136].

Recall that an i.i.d. sequence $(Y_i)_{i=1}^{\infty}$ of random variables is in the *domain of attraction* of a random variable X if there exist sequences of constants $(A_i)_{i=1}^{\infty}$ and $(B_i)_{i=1}^{\infty}$ so that

$$B_n \sum_{i=1}^{n} Y_i - A_n$$

converges in distribution to X as $n \to \infty$. $(Y_i)_{i=1}^{\infty}$ is in the *domain of normal attraction* of X if

$$B_n = \frac{1}{bn^{1/\alpha}}$$

for some $b > 0$.

We confine our search to the direction suggested by (iii) because of the particular importance classically of results of this type (cf. the central limit theorem). Our main findings are summarized in Theorem III.11, which details sufficient conditions for membership in the domain of normal attraction of a given α-stable random variable. While the initial appearance of our distributional hypotheses is perhaps forbidding, in point of fact, our assumptions are manageable. This is illustrated by the discussion immediately following Theorem III.11, as well as Examples III.13 and III.14.

Example III.13 establishes that the α-stable random variables under consideration are in their own domain of normal attraction. Although one need not apply Theorem III.11 for this purpose, the writeup serves a clarifying role and any credible result clearly must pass this litmus test.

Example III.14 is more substantive. Setting aside a few mild "uniformity" conditions which arise due to the supremum, this example can be understood in an intuitive manner (see Section 3.4). This falls out of our analysis just below Theorem III.11, where we describe the relationship between our work and the classical result noted in (iii) above. More specifically, Theorem III.11 detects all classical random variables in this collection with mean zero and a cumulative distribution function (cdf) that satisfies a small differentiability requirement. An extra regularity condition on the cdf is unavoidable, as one must translate its form into properties that can be stated only in terms of expectation.

The strategy of our proof is to reduce demonstrating convergence in distribution to showing that a certain limit involving the solution to the backward version of our generating PIDE is zero. Upon breaking up our domain and summing the corresponding increments of the solution, regularity properties of this function are employed to argue that size of the terms being added together decay rapidly enough in the limit to furnish the desired conclusion. This general scheme is similar to that initiated in [190], except that the generating equation there is

$$\partial_t u - \frac{1}{2}\left(\overline{\sigma}^2 \left(\partial_{xx} u\right)^+ - \underline{\sigma}^2 \left(\partial_{xx} u\right)^-\right) = 0, \quad (t, x) \in (0, \infty) \times \mathbb{R}$$
$$u(0, x) = \psi(x), \quad x \in \mathbb{R}$$

for some $0 \leq \underline{\sigma}^2 \leq \overline{\sigma}^2$ and appropriate function ψ. Recall that this equation is known as the Barenblatt equation if $\underline{\sigma}^2 > 0$ and has been studied in [37] and [28], for instance. Ours is given by (3.1), a difference that leads to a few difficulties as reflected by the increased complexity of our hypotheses. To overcome these difficulties, we use the technology from [157], [156], and [64].

The work in this paper offers a step toward understanding α-stability under sublinear expectation. The simple interpretation admitted by Example III.14 is promising, as developing intuition in this environment is usually a tough undertaking for the reasons mentioned previously.

A brief overview of necessary background material can be found in Section 3.2. We prove our main result and discuss its connection to the classical case in Section 3.3. Examples highlighting the applications of our main result are contained in Section 3.4. We give some prerequisite material for the proof of the essential interior regularity estimate for our PIDE in Section 3.5. The proof of this estimate is in Section 3.6.

3.2 Background

We now offer a concise account of those aspects of sublinear expectations, α-stable random variables, and PIDEs which are required for the sequel.[1] References for more comprehensive treatments of these topics are also included for the convenience of the

[1] Further information on PIDE interior regularity theory is contained in Section 3.5.

interested reader.

Definition III.1. Let $\mathcal{H}$ be a collection of real-valued functions on a set Ω. A *sublinear expectation* is an operator $\mathcal{E} : \mathcal{H} \longrightarrow \mathbb{R}$ which is

(i) monotonic: $\mathcal{E}[X] \leq \mathcal{E}[Y]$ if $X \leq Y$,

(ii) constant-preserving: $\mathcal{E}[c] = c$ for any $c \in \mathbb{R}$,

(iii) sub-additive: $\mathcal{E}[X+Y] \leq \mathcal{E}[X] + \mathcal{E}[Y]$, and

(iv) positive homogeneous: $\mathcal{E}[\lambda X] = \lambda \mathcal{E}[X]$ for $\lambda \geq 0$.

The triple $(\Omega, \mathcal{H}, \mathcal{E})$ is called a *sublinear expectation space.*

One views $\mathcal{H}$ as a space of random variables on Ω. Typically, it is assumed that $\mathcal{H}$

(i) is a linear space,

(ii) contains all constant functions, and

(iii) contains $\psi(X_1, X_2, \ldots, X_n)$ for every $X_1, X_2, \ldots, X_n \in \mathcal{H}$ and $\psi \in C_{b.Lip}(\mathbb{R}^n)$, where $C_{b.Lip}(\mathbb{R}^n)$ is the set of bounded Lipschitz functions on $\mathbb{R}^n$;

however, we will expend little attention on either Ω or $\mathcal{H}$. Delicacy needs to be exercised while computing sublinear expectations. A rare instance when a classical technique can be justly employed is the following.

Lemma III.2. *Consider two random variables $X, Y \in \mathcal{H}$ such that $\mathcal{E}[Y] = -\mathcal{E}[-Y]$. Then*

$$\mathcal{E}[X + \alpha Y] = \mathcal{E}[X] + \alpha \mathcal{E}[Y]$$

for all $\alpha \in \mathbb{R}$.

This result is notably useful in the case where $\mathcal{E}[Y] = \mathcal{E}[-Y] = 0$.

Definition III.3. A random variable $Y \in \mathcal{H}$ is said to be *independent* from a random variable $X \in \mathcal{H}$ if for all $\psi \in C_{b.Lip}(\mathbb{R}^2)$, we have

$$\mathcal{E}[\psi(X,Y)] = \mathcal{E}\left[\mathcal{E}[\psi(x,Y)]_{x=X}\right].$$

Observe the deliberate wording. This choice is crucial, as independence can be asymmetric in our context. Note that this definition reduces to the traditional one if $\mathcal{E}$ is a classical expectation. The same is true for the next three concepts.

Definition III.4. Let X, Y, and $(Y_n)_{n=1}^{\infty}$ be random variables, i.e., X, Y, and $(Y_n)_{n=1}^{\infty} \in \mathcal{H}$.

(i) X and Y are *identically distributed*, denoted $X \sim Y$, if

$$\mathcal{E}[\psi(X)] = \mathcal{E}[\psi(Y)]$$

for all $\psi \in C_{b.Lip}(\mathbb{R})$.

(ii) If X and Y are identically distributed and Y is independent from X, then Y is an *independent copy* of X.

(iii) $(Y_n)_{n=1}^{\infty}$ *converges in distribution* to Y, which we denote by $Y_n \xrightarrow{d} Y$, if

$$\lim_{n\to\infty} \mathcal{E}[\psi(Y_n)] = \mathcal{E}[\psi(Y)]$$

for all $\psi \in C_{b.Lip}(\mathbb{R})$.

Random variables need not be defined on the same space to have appropriate notions of (i) or (iii). In this case, the above definitions require the obvious notational modifications. Further details concerning general sublinear expectation spaces can be found in [188] or [194].

Definition III.5. Let $\alpha \in (0,2]$. A random variable X is said to be *(strictly) α-stable* if for all a, $b \geq 0$,

$$aX + bY$$

and

$$\left(a^\alpha + b^\alpha\right)^{1/\alpha} X$$

are identically distributed, where Y is an independent copy of X.

Three examples of α-stable random variables exist in the current literature. For $\alpha = 1$, there are the maximal random variables discussed in references such as [191], [194], and [128]. When $\alpha = 2$, we have the G-normal random variables of Peng. Resources on this topic are plentiful and include [188], [193], [194], [163], and [47]. If $\alpha \in (1,2)$, we can consider X_1 for a nonlinear α-stable Lévy process $(X_t)_{t\geq 0}$ in the framework of [173]. Our focus shall be restricted to the last situation.

The construction of nonlinear Lévy processes in [173] extends that studied in [129], [208], [186], and [185] and is much more general than our present objectives demand. We limit our presentation to a few key ideas. Let

(i) $\alpha \in (1,2)$;

(ii) $K_\pm$ be a bounded measurable subset of $\mathbb{R}_+$;

(iii) $F_{k_\pm}$ be the α-stable Lévy measure

$$F_{k_\pm}(dz) = \left(k_- \mathbf{1}_{(-\infty,0)} + k_+ \mathbf{1}_{(0,\infty)}\right)(z)\,|z|^{-\alpha-1} dz$$

for all $k_\pm \in K_\pm$; and

(iv) $\Theta = \left\{\left(0, 0, F_{k_\pm}\right) : k_\pm \in K_\pm\right\}$.

One can then produce a process $(X_t)_{t\geq 0}$ which is a nonlinear Lévy process whose local characteristics are described by the set of Lévy triplets Θ. This means the following.

(i) $(X_t)_{t\geq 0}$ is a real-valued càdlàg process.

(ii) $X_0 = 0$.

(iii) $(X_t)_{t\geq 0}$ has stationary increments, i.e., $X_t - X_s$ and X_{t-s} are identically distributed for all $0 \leq s \leq t$.

(iv) $(X_t)_{t\geq 0}$ has independent increments, i.e., $X_t - X_s$ is independent from $(X_{s_1}, \ldots, X_{s_n})$ for all $0 \leq s_1 \leq \cdots \leq s_n \leq s \leq t$.

(v) If $\psi \in C_{b.Lip}(\mathbb{R})$ and u is defined by

$$u(t,x) = \mathcal{E}\left[\psi(x + X_t)\right]$$

for all $(t,x) \in [0,\infty) \times \mathbb{R}$, then u is the unique[2] viscosity solution[3] of

$$\partial_t u(t,x) - \sup_{k_\pm \in K_\pm} \left\{ \int_{\mathbb{R}} \delta_z u(t,x)\, F_{k_\pm}(dz) \right\} = 0, \quad (t,x) \in (0,\infty) \times \mathbb{R}$$
$$u(0,x) = \psi(x), \quad x \in \mathbb{R}. \tag{3.1}$$

[2]The uniqueness of a viscosity solution of (3.1) can be viewed as a special case of Theorem 2.5 in [173].

[3]We take the following definition from Section 2.2 of [173]. Let $C_b^{2,3}((0,\infty) \times \mathbb{R})$ denote the set of functions on $(0,\infty) \times \mathbb{R}$ having bounded continuous partial derivatives up to the second and third order in t and x, respectively. A bounded upper semicontinuous function u on $[0,\infty) \times \mathbb{R}$ is a *viscosity subsolution* of (3.1) if

$$u(0,\cdot) \leq \psi(\cdot)$$

and for any $(t,x) \in (0,\infty) \times \mathbb{R}$,

$$\partial_t \varphi(t,x) - \sup_{k_\pm \in K_\pm} \left\{ \int_{\mathbb{R}} \delta_z \varphi(t,x)\, F_{k_\pm}(dz) \right\} \leq 0$$

whenever $\varphi \in C_b^{2,3}((0,\infty) \times \mathbb{R})$ is such that

$$\varphi \geq u$$

on $(0,\infty) \times \mathbb{R}$ and

$$\varphi(t,x) = u(t,x).$$

To define a *viscosity supersolution* of (3.1), one reverses the inequalities and semicontinuity. A bounded continuous function is a *viscosity solution* of (3.1) if it is both a viscosity subsolution and supersolution. Viscosity solutions of the other PIDEs appearing in this paper, e.g., see Lemma III.18, are defined similarly.

Here we use the notation

$$\delta_z u(t,x) := u(t,x+z) - u(t,x) - \partial_x u(t,x)\, z$$

since the right hand side of this equation as well as similar expressions will frequently occur throughout the paper.

A critical feature of this setup is that if Θ is a singleton, $(X_t)_{t\geq 0}$ is a classical Lévy process with triplet Θ. That X_1 actually is an α-stable random variable is not immediately obvious. We give a brief argument in Example III.13, but the core of this observation is a result from [173] (see Example 2.7).

Lemma III.6. *For all $\beta > 0$ and $t \geq 0$, $X_{\beta t}$ and $\beta^{1/\alpha} X_t$ are identically distributed.*

The dynamic programming principle in Lemma III.7 (see Lemma 5.1 in [173]) and the absolute value bound in Lemma III.8 (see Lemma 5.2 in [173]) also play a central role when using our main result to check that X_1 is in its own domain of normal attraction.

Lemma III.7. *For all $0 \leq s \leq t < \infty$ and $x \in \mathbb{R}$,*

$$u(t,x) = \mathcal{E}\left[u(t-s, x+X_s)\right].$$

Lemma III.8. *We have that*

$$\mathcal{E}\left[|X_1|\right] < \infty.$$

The remaining essential ingredients for our purposes describe the regularity of u. The first result describes properties of u which are valid on the whole domain. It is a special case of Lemma 5.3 in [173].

Lemma III.9. *The function u is uniformly bounded by $\|\psi\|_{L^\infty(\mathbb{R})}$ and jointly continuous. More precisely, $u(t,\cdot)$ is Lipschitz continuous with constant Lip(ψ), the*

Lipschitz constant of ψ, and $u(\cdot, x)$ is locally 1/2-Hölder continuous with a constant depending only on Lip(ψ) and

$$\sup_{k_\pm \in K_\pm} \left\{ \int_{\mathbb{R}} |z| \wedge |z|^2 \, F_{k_\pm}(dz) \right\} < \infty.$$

We will require even stronger regularity estimates for u. To obtain these, we must restrict our attention to the interior of the domain.

Proposition III.10. *Suppose that for some λ, $\Lambda > 0$, we know $\lambda < k_\pm < \Lambda$ for all $k_\pm \in K_\pm$. For any $h > 0$,*

(i) $\partial_t u$ and $\partial_x u$ exist and are bounded on $[h, h+1] \times \mathbb{R}$;

(ii) there are constants C, $\gamma > 0$ such that

$$|\partial_t u(t_0, x) - \partial_t u(t_1, x)| \le C |t_0 - t_1|^{\gamma/\alpha}$$

$$|\partial_t u(t, x_0) - \partial_t u(t, x_1)| \le C |x_0 - x_1|^{\gamma}$$

for all (t_0, x), (t_1, x), (t, x_0), $(t, x_1) \in [h, h+1] \times \mathbb{R}$;

(iii) u is a classical solution of (3.1) on $[h, h+1] \times \mathbb{R}$; and

(iv) if $K_\pm$ contains exactly one pair $\{k_\pm\}$, then $\partial^2_{xx} u$ exists and is bounded on $[h, h+1] \times \mathbb{R}$.

The proof of Proposition III.10 can be found in Section 3.6.

3.3 Main Result

To facilitate our discussion in the sequel, we now fix some notation. Compared with Section 3.2, we make only one alteration to our nonlinear α-stable Lévy process $(X_t)_{t \ge 0}$: additionally assume that $K_\pm$ is a subset of (λ, Λ) for some λ, $\Lambda > 0$. We will make use of this in conjunction with Proposition III.10.

We also consider a sequence $(Y_i)_{i=1}^{\infty}$ of random variables on some sublinear expectation space. The only aspect of this space that we will invoke directly is the sublinear expectation itself, say $\mathcal{E}'$. Distinguishing between $\mathcal{E}$ and $\mathcal{E}'$ will be convenient for Example III.14. We further specify that $(Y_i)_{i=1}^{\infty}$ is i.i.d. in the sense that Y_{i+1} is independent from $(Y_1, Y_2, \ldots, Y_i)$ and $Y_{i+1} \sim Y_i$ for all $i \geq 1$. After proper normalization,

$$S_n := \sum_{i=1}^{n} Y_i$$

will be the sequence attracted to X_1.

Theorem III.11. *Suppose that*

(i) $\mathcal{E}'[Y_1] = \mathcal{E}'[-Y_1] = 0$*;*

(ii) $\mathcal{E}'[|Y_1|] < \infty$*; and*

(iii) for any $0 < h < 1$ *and* $\psi \in C_{b.Lip}(\mathbb{R})$*,*

$$n\left|\mathcal{E}'\left[\delta_{B_nY_1} v(t,x)\right] - \left(\frac{1}{n}\right) \sup_{k_\pm \in K_\pm} \left\{\int_{\mathbb{R}} \delta_z v(t,x)\, F_{k_\pm}(dz)\right\}\right| \to 0 \qquad (3.2)$$

uniformly on $[0,1] \times \mathbb{R}$ *as* $n \to \infty$*, where* v *is the unique viscosity solution of*

$$\partial_t v(t,x) + \sup_{k_\pm \in K_\pm} \left\{\int_{\mathbb{R}} \delta_z v(t,x)\, F_{k_\pm}(dz)\right\} = 0, \quad (t,x) \in (-h, 1+h) \times \mathbb{R}$$

$$v(1+h, x) = \psi(x), \quad x \in \mathbb{R}. \qquad (3.3)$$

Then

$$B_n S_n \xrightarrow{d} X_1$$

as $n \to \infty$*.*

Admittedly, a cursory glance over our hypotheses leaves one with the impression that they are intractable. The opposite is true. Before presenting the proof of

Theorem III.11, let us demonstrate that when our attention is confined to the classical case, we are imposing only a mild and natural supplementary restriction on the attracted random variable. In addition to being a significant remark in itself, this work also underlies Example III.14.

Assume that Θ is a singleton. Since $(X_t)_{t\geq 0}$ is the classical Lévy process with triplet $\left(0, 0, F_{k_\pm}\right)$, the characteristic function of X_1, denoted φ_{X_1}, is given by

$$\varphi_{X_1}(t) = \exp\left(k_- \int_{-\infty}^{0} \frac{\exp(itz) - 1 - itz}{|z|^{\alpha+1}}\, dz + k_+ \int_{0}^{\infty} \frac{\exp(itz) - 1 - itz}{z^{\alpha+1}}\, dz\right)$$

for all $t \in \mathbb{R}$. In the case where Y_1 is a classical random variable with mean zero, Theorem 2.6.7 from [136] implies that

$$B_n S_n \xrightarrow{d} X_1$$

as $n \to \infty$ if and only if the cdf of Y_1, denoted F_{Y_1}, has the form

$$F_{Y_1}(z) = \begin{cases} \left[b^\alpha (k_-/\alpha) + \beta_1(z)\right] \frac{1}{|z|^\alpha} & z < 0 \\ 1 - \left[b^\alpha (k_+/\alpha) + \beta_2(z)\right] \frac{1}{z^\alpha} & z > 0 \end{cases}$$

for some functions β_1 and β_2 satisfying

$$\lim_{z\to-\infty} \beta_1(z) = \lim_{z\to\infty} \beta_2(z) = 0.$$

As there is no appropriate counterpart of the cdf in the sublinear setting, we must recast this condition using expectation. To do so requires F_{Y_1} to possess further regularity properties. For convenience, say that after an extension, the β_i's are continuously differentiable on their respective closed half-lines. This is the lone extra requirement we shall need.

It follows that

$$\mathbb{E}\left[|Y_1|\right] < \infty$$

since

$$\int_0^\infty z\,dF_{Y_1}(z) = -\int_0^1 \frac{\beta_2'(z)}{z^{\alpha-1}}\,dz + \int_0^1 \frac{b^\alpha k_+ + \alpha\beta_2(z)}{z^\alpha}\,dz + \beta_2(1)$$
$$+ \int_1^\infty \frac{\beta_2(z)}{z^\alpha}\,dz + \int_1^\infty \frac{b^\alpha k_+}{z^\alpha}\,dz$$
$$< \infty \tag{3.4}$$

and similarly for the integral along the negative half-line. One could have cited Theorem 2.6.4 of [136] instead, but (3.4) will be helpful in Example III.14. We also get

$$n\left|\mathbb{E}\left[\delta_{B_nY_1}v(t,x)\right] - \left(\frac{1}{n}\right)\int_{\mathbb{R}} \delta_z v(t,x)\,F_{k_\pm}(dz)\right|$$
$$= \left(\frac{1}{b^\alpha}\right)\left|\int_{\mathbb{R}} \delta_z v(t,x)\left(\frac{\beta_1'(B_n^{-1}z)\,|B_n^{-1}z| + \alpha\beta_1(B_n^{-1}z)}{|z|^{\alpha+1}}\mathbf{1}_{(-\infty,0)}(z)\right.\right.$$
$$\left.\left.+\frac{-\beta_2'(B_n^{-1}z)\,|B_n^{-1}z| + \alpha\beta_2(B_n^{-1}z)}{|z|^{\alpha+1}}\mathbf{1}_{(0,\infty)}(z)\right)dz\right| \tag{3.5}$$

for all $(t,x) \in [0,1]\times\mathbb{R}$ and $n \geq 1$ by changing variables.

A careful application of elementary estimates shows that this last expression tends to zero uniformly on $[0,1]\times\mathbb{R}$ as $n\to\infty$. To see this, note that we can choose an upper bound, say M_1, for $|\partial_{xx}v|$, $|\partial_x v|$, and $|v|$ on $[0,1]\times\mathbb{R}$ by Lemma III.9 and Proposition III.10. Then using integration by parts and the dominated convergence theorem,

$$\left|\int_1^\infty \delta_z v(t,x)\left(\frac{-\beta_2'(B_n^{-1}z)\,|B_n^{-1}z| + \alpha\beta_2(B_n^{-1}z)}{|z|^{\alpha+1}}\right)dz\right|$$
$$= \left|\delta_1 v(t,x)\,\beta_2\left(B_n^{-1}\right) + \int_1^\infty \frac{\beta_2(B_n^{-1}z)}{z^\alpha}\left[\partial_x v(t,x+z) - \partial_x v(t,x)\right]\,dz\right|$$
$$\leq 3M_1\left|\beta_2\left(B_n^{-1}\right)\right| + 2M_1\int_1^\infty \frac{|\beta_2(B_n^{-1}z)|}{z^\alpha}\,dz$$
$$\to 0 \tag{3.6}$$

as $n \to \infty$. The mean value theorem and a change of variables give

$$\begin{aligned}
&\left| \int_0^{B_n} \delta_z v(t,x) \left(\frac{-\beta_2'(B_n^{-1}z)\,|B_n^{-1}z| + \alpha\beta_2(B_n^{-1}z)}{|z|^{\alpha+1}} \right) dz \right| \\
&\le \int_0^{B_n} M_1 \frac{|-\beta_2'(B_n^{-1}z)(B_n^{-1}z) + \alpha\beta_2(B_n^{-1}z)|}{z^{\alpha-1}}\, dz \\
&= \left(\frac{M_1}{b^{2-\alpha} n^{\frac{2}{\alpha}-1}} \right) \int_0^1 \frac{|-\beta_2'(z)\,z + \alpha\beta_2(z)|}{z^{\alpha-1}}\, dz \\
&\to 0 \qquad (3.7)
\end{aligned}$$

as $n \to \infty$. We have

$$\begin{aligned}
&\left| \int_{B_n}^1 \delta_z v(t,x) \left(\frac{\alpha\beta_2(B_n^{-1}z)}{|z|^{\alpha+1}} \right) dz \right| \\
&\le \int_{B_n}^1 M_1 \frac{|\alpha\beta_2(B_n^{-1}z)|}{z^{\alpha-1}}\, dz \\
&\le M_1\alpha \int_0^1 \frac{|\beta_2(B_n^{-1}z)|}{z^{\alpha-1}}\, dz \\
&\to 0 \qquad (3.8)
\end{aligned}$$

as $n \to \infty$ by the mean value theorem and dominated convergence theorem. Finally,

$$\begin{aligned}
&\left| \int_{B_n}^1 \delta_z v(t,x) \left(\frac{-\beta_2'(B_n^{-1}z)(B_n^{-1}z)}{|z|^{\alpha+1}} \right) dz \right| \\
&= \Bigg| -\delta_1 v(t,x)\,\beta_2\left(B_n^{-1}\right) + \delta_{B_n} v(t,x)\,(B_n)^{-\alpha}\,\beta_2(1) \\
&\quad + \int_{B_n}^1 \left[\partial_x v(t,x+z) - \partial_x v(t,x)\right] \left(\frac{\beta_2(B_n^{-1}z)}{z^{\alpha}} \right) dz \\
&\quad - \alpha \int_{B_n}^1 \delta_z v(t,x) \left(\frac{\beta_2(B_n^{-1}z)}{z^{\alpha+1}} \right) dz \Bigg| \\
&\le 3M_1 \left|\beta_2\left(B_n^{-1}\right)\right| + M_1\,|\beta_2(1)| \left(\frac{1}{b^{2-\alpha} n^{\frac{2}{\alpha}-1}} \right) + \int_{B_n}^1 M_1 \frac{|\beta_2(B_n^{-1}z)|}{z^{\alpha-1}} dz \\
&\quad + \alpha \int_{B_n}^1 M_1 \frac{|\beta_2(B_n^{-1}z)|}{z^{\alpha-1}} dz \\
&\le 3M_1 \left|\beta_2\left(B_n^{-1}\right)\right| + M_1\,|\beta_2(1)| \left(\frac{1}{b^{2-\alpha} n^{\frac{2}{\alpha}-1}} \right) + 2\alpha M_1 \int_0^1 \frac{|\beta_2(B_n^{-1}z)|}{z^{\alpha-1}}\, dz \\
&\to 0 \qquad (3.9)
\end{aligned}$$

as $n \to \infty$ by integration by parts, the dominated convergence theorem, and the mean value theorem. The integrals along the negative half-line are handled similarly.

Having established the connection between Theorem III.11 and the classical case, we finally present its proof.

Proof of Theorem III.11. We need to show that

$$\lim_{n\to\infty} \mathcal{E}' \left[\psi\left(B_n S_n\right)\right] = \mathcal{E}\left[\psi\left(X_1\right)\right] \tag{3.10}$$

for all $\psi \in C_{b.Lip}(\mathbb{R})$. Our initial step will be to reduce proving (3.10) to proving (3.13). The advantage of doing so is that we can then incorporate the regularity properties described in Lemma III.9 and Proposition III.10. These properties alone do much of the heavy lifting in the estimates at the heart of the argument, and our distributional assumptions do the rest.

Let $\psi \in C_{b.Lip}(\mathbb{R})$, and define u by

$$u\left(t,x\right) = \mathcal{E}\left[\psi\left(x + X_t\right)\right] \tag{3.11}$$

for all $(t,x) \in [0,\infty) \times \mathbb{R}$. We know from Section 3.2 that u is the unique viscosity solution of (3.1).

It will be more convenient for our purposes to work with the backward equation. Since we will soon rely on the interior regularity results of Proposition III.10, we also let $0 < h < 1$ and define v by

$$v\left(t,x\right) = u\left(1+h-t,x\right) \tag{3.12}$$

for $(t,x) \in (-h, 1+h] \times \mathbb{R}$. Then v will be the unique viscosity solution of (3.3).

Observe that v inherits key regularity properties from u. At the moment, it is enough to note that for any $(t,x) \in (-h,1+h] \times \mathbb{R}$, $v(\cdot,x)$ is 1/2-Hölder continuous with some constant K_1 and $v(t,\cdot)$ is Lipschitz continuous with constant $\mathrm{Lip}(\psi)$ by

Lemma III.9. Because the t-domain has length $1+2h$ and $0<h<1$, the 1/2-Hölder continuity is uniform, and we can assume that K_1 does not depend on h. It follows by (3.11) and (3.12) that

$$
\begin{aligned}
&\limsup_{n\to\infty} \left|\mathcal{E}'\left[\psi\left(B_nS_n\right)\right]-\mathcal{E}\left[\psi\left(X_1\right)\right]\right| \\
&\leq \limsup_{n\to\infty} \left(\left|\mathcal{E}'\left[\psi\left(B_nS_n\right)\right]-\mathcal{E}'\left[v\left(1,B_nS_n\right)\right]\right|+\left|\mathcal{E}'\left[v\left(1,B_nS_n\right)\right]-v\left(0,0\right)\right|\right. \\
&\qquad \left.+\left|v\left(0,0\right)-\mathcal{E}\left[\psi\left(X_1\right)\right]\right|\right) \\
&= \limsup_{n\to\infty} \left(\left|\mathcal{E}'\left[v\left(1+h,B_nS_n\right)\right]-\mathcal{E}'\left[v\left(1,B_nS_n\right)\right]\right|+\left|\mathcal{E}'\left[v\left(1,B_nS_n\right)\right]-v\left(0,0\right)\right|\right. \\
&\qquad \left.+\left|v\left(0,0\right)-v\left(h,0\right)\right|\right) \\
&\leq \limsup_{n\to\infty} \left(\mathcal{E}'\left[K_1\sqrt{h}\right]+\left|\mathcal{E}'\left[v\left(1,B_nS_n\right)\right]-v\left(0,0\right)\right|\right)+K_1\sqrt{h} \\
&= 2K_1\sqrt{h}+\limsup_{n\to\infty} \left|\mathcal{E}'\left[v\left(1,B_nS_n\right)\right]-v\left(0,0\right)\right|.
\end{aligned}
$$

As h is arbitrary, it is sufficient to show that

$$
\lim_{n\to\infty} \mathcal{E}'\left[v\left(1,B_nS_n\right)\right]=v\left(0,0\right). \tag{3.13}
$$

The required estimates are intricate, so we will give them in Lemma III.12 below. □

Lemma III.12. *In the setup of Theorem III.11,*

$$
\lim_{n\to\infty} \mathcal{E}'\left[v\left(1,B_nS_n\right)\right]=v\left(0,0\right).
$$

Proof of Lemma III.12. For all $n\geq 3$,

$$
\begin{aligned}
&v\left(1,B_nS_n\right)-v\left(0,0\right) \\
&\quad = v\left(1,B_nS_n\right)-v\left(\frac{n-1}{n},B_nS_n\right)+\sum_{i=2}^{n-1}\left[v\left(\frac{i}{n},B_nS_{i+1}\right)-v\left(\frac{i-1}{n},B_nS_i\right)\right] \\
&\qquad + v\left(\frac{1}{n},B_nS_2\right)-v\left(0,0\right).
\end{aligned} \tag{3.14}
$$

Our analysis now becomes delicate. We would like to show that when we apply $\mathcal{E}'$ to (3.14) and let $n \to \infty$, the result goes to zero. Since the number of terms in this decomposition is growing with n, we must prove that our v-increments are decaying quite rapidly. The properties of v arising from Lemma III.9 are only enough to manage the first and last terms. By the 1/2-Hölder continuity of $v(\cdot, x)$,

$$\mathcal{E}'\left[\left|v(1, B_n S_n) - v\left(\frac{n-1}{n}, B_n S_n\right)\right|\right] \leq \mathcal{E}'\left[K_1\sqrt{\frac{1}{n}}\right] = K_1\sqrt{\frac{1}{n}}. \tag{3.15}$$

If we also use the Lipschitz continuity of $v(t, \cdot)$ and the fact that Y_2 is independent from Y_1, we get

$$\begin{aligned}
&\mathcal{E}'\left[\left|v\left(\frac{1}{n}, B_n S_2\right) - v(0,0)\right|\right] \\
&\leq \mathcal{E}'\left[\left|v\left(\frac{1}{n}, B_n S_2\right) - v(0, B_n S_2)\right|\right] + \mathcal{E}'\left[\left|v(0, B_n S_2) - v(0,0)\right|\right] \\
&\leq \mathcal{E}'\left[K_1\sqrt{\frac{1}{n}}\right] + \mathcal{E}'\left[\mathrm{Lip}(\psi) B_n |S_2|\right] \\
&\leq K_1\sqrt{\frac{1}{n}} + 2\mathrm{Lip}(\psi) B_n \mathcal{E}'\left[|Y_1|\right].
\end{aligned} \tag{3.16}$$

We remark that although we only referred to $C_{b.Lip}(\mathbb{R})$ in our definition of independence, our manipulations are still valid by Exercise 3.20 in [194].

Proposition III.10 allows us to control the remaining terms. Again, this motivates our requirement that $K_\pm \subset (\lambda, \Lambda)$ for some $0 < \lambda < \Lambda$. We can find a constant $K_2 > 0$ such that $\partial_t v$ exists on $[0,1] \times \mathbb{R}$ and

$$\begin{aligned}
\left|\partial_t v(t_0, x) - \partial_t v(t_1, x)\right| &\leq K_2 \left|t_0 - t_1\right|^{\gamma/\alpha} \\
\left|\partial_t v(t, x_0) - \partial_t v(t, x_1)\right| &\leq K_2 \left|x_0 - x_1\right|^{\gamma}
\end{aligned} \tag{3.17}$$

for all (t_0, x), (t_1, x), (t, x_0), and $(t, x_1) \in [0,1] \times \mathbb{R}$. We then break down the rest of

(3.14) a bit further. If $2 \le i \le n-1$,

$$\begin{aligned}
&v\left(\frac{i}{n}, B_n S_{i+1}\right) - v\left(\frac{i-1}{n}, B_n S_i\right) \\
&= v\left(\frac{i}{n}, B_n S_{i+1}\right) - v\left(\frac{i-1}{n}, B_n S_{i+1}\right) - \partial_t v\left(\frac{i-1}{n}, B_n S_i\right)\frac{1}{n} \\
&\quad + \partial_t v\left(\frac{i-1}{n}, B_n S_i\right)\frac{1}{n} + v\left(\frac{i-1}{n}, B_n S_{i+1}\right) - v\left(\frac{i-1}{n}, B_n S_i\right).
\end{aligned}$$

Let

$$C_i^n = v\left(\frac{i}{n}, B_n S_{i+1}\right) - v\left(\frac{i-1}{n}, B_n S_{i+1}\right) - \partial_t v\left(\frac{i-1}{n}, B_n S_i\right)\frac{1}{n}$$

and

$$D_i^n = \partial_t v\left(\frac{i-1}{n}, B_n S_i\right)\frac{1}{n} + v\left(\frac{i-1}{n}, B_n S_{i+1}\right) - v\left(\frac{i-1}{n}, B_n S_i\right).$$

We can establish an appropriate bound for the C_i^n's using (3.17):

$$\begin{aligned}
|C_i^n| &= \left| \frac{1}{n}\int_0^1 \left[\partial_t v\left(\frac{i-1+\beta}{n}, B_n S_{i+1}\right) - \partial_t v\left(\frac{i-1}{n}, B_n S_{i+1}\right)\right] d\beta \right. \\
&\qquad \left. + \frac{1}{n}\left[\partial_t v\left(\frac{i-1}{n}, B_n S_{i+1}\right) - \partial_t v\left(\frac{i-1}{n}, B_n S_i\right)\right]\right| \\
&\le \frac{1}{n}\int_0^1 \left|\partial_t v\left(\frac{i-1+\beta}{n}, B_n S_{i+1}\right) - \partial_t v\left(\frac{i-1}{n}, B_n S_{i+1}\right)\right| d\beta \\
&\qquad + \frac{1}{n}\left|\partial_t v\left(\frac{i-1}{n}, B_n S_{i+1}\right) - \partial_t v\left(\frac{i-1}{n}, B_n S_i\right)\right| \\
&\le \frac{1}{n}\int_0^1 K_2 \left|\frac{\beta}{n}\right|^{\gamma/\alpha} d\beta + \frac{1}{n} K_2 B_n^\gamma |Y_{i+1}|^\gamma \\
&\le \frac{K_2}{n}\left[\left(\frac{1}{n}\right)^{\gamma/\alpha} + B_n^\gamma |Y_{i+1}|^\gamma\right].
\end{aligned}$$

Hence, for $2 \le i \le n-1$,

$$\mathcal{E}'\left[|C_i^n|\right] \le \frac{K_2}{n}\left[\left(\frac{1}{n}\right)^{\gamma/\alpha} + B_n^\gamma \mathcal{E}'\left[|Y_1|^\gamma\right]\right] \tag{3.18}$$

since Y_{i+1} and Y_1 are identically distributed. Note that hypothesis (ii) gives that

$$\mathcal{E}'\left[|Y_1|^\gamma\right] < \infty.$$

While we need (3.17) to bound the D_i^n's, we finally use (3.2) too. Let $\epsilon > 0$. By (3.2), we can find N such that $n \geq N$ implies

$$n\left|\mathcal{E}'\left[\delta_{B_nY_1}v\left(t,x\right)\right]-\left(\frac{1}{n}\right)\sup_{k_\pm\in K_\pm}\left\{\int_{\mathbb{R}}\delta_z v\left(t,x\right)F_{k_\pm}\left(dz\right)\right\}\right|<\epsilon$$

on $[0,1]\times\mathbb{R}$. Now

$$\mathcal{E}'\left[v\left(\frac{i-1}{n},B_nx+B_nY_1\right)\right]-v\left(\frac{i-1}{n},B_nx\right)=\mathcal{E}'\left[\delta_{B_nY_1}v\left(\frac{i-1}{n},B_nx\right)\right]$$

by (i), so for these n,

$$\begin{aligned}
&n\left|v\left(\frac{i-2}{n},B_nx\right)-\mathcal{E}'\left[v\left(\frac{i-1}{n},B_nx+B_nY_1\right)\right]\right|\\
&\quad=n\Bigg|v\left(\frac{i-2}{n},B_nx\right)-\mathcal{E}'\left[v\left(\frac{i-1}{n},B_nx+B_nY_1\right)\right]\\
&\qquad+v\left(\frac{i-1}{n},B_nx\right)-v\left(\frac{i-1}{n},B_nx\right)\\
&\qquad+\left(\frac{1}{n}\right)\partial_t v\left(\frac{i-1}{n},B_nx\right)+\left(\frac{1}{n}\right)\sup_{k_\pm\in K_\pm}\left\{\int_{\mathbb{R}}\delta_z v\left(\frac{i-1}{n},B_nx\right)F_{k_\pm}\left(dz\right)\right\}\Bigg|\\
&\quad\leq\left|-\frac{v\left(\frac{i-2}{n},B_nx\right)-v\left(\frac{i-1}{n},B_nx\right)}{-1/n}+\partial_t v\left(\frac{i-1}{n},B_nx\right)\right|\\
&\qquad+n\left|\mathcal{E}'\left[\delta_{B_nY_1}v\left(\frac{i-1}{n},B_nx\right)\right]-\left(\frac{1}{n}\right)\sup_{k_\pm\in K_\pm}\left\{\int_{\mathbb{R}}\delta_z v\left(\frac{i-1}{n},B_nx\right)F_{k_\pm}\left(dz\right)\right\}\right|\\
&\quad<\frac{K_2}{n^{\gamma/\alpha}}+\epsilon
\end{aligned}$$

by the mean value theorem, (3.3), and (3.17). Then

$$\begin{aligned}
&\left|\partial_t v\left(\frac{i-1}{n},B_nx\right)\frac{1}{n}+\mathcal{E}'\left[v\left(\frac{i-1}{n},B_nx+B_nY_{i+1}\right)\right]-v\left(\frac{i-1}{n},B_nx\right)\right|\\
&\leq\frac{1}{n}\left|\partial_t v\left(\frac{i-1}{n},B_nx\right)+\frac{v\left(\frac{i-2}{n},B_nx\right)-v\left(\frac{i-1}{n},B_nx\right)}{1/n}\right|\\
&\qquad+\left|\mathcal{E}'\left[v\left(\frac{i-1}{n},B_nx+B_nY_1\right)\right]-v\left(\frac{i-2}{n},B_nx\right)\right|\\
&<\frac{2K_2}{n^{1+\gamma/\alpha}}+\frac{\epsilon}{n}
\end{aligned}\tag{3.19}$$

for $2 \leq i \leq n-1$, $x \in \mathbb{R}$, and $n \geq N$.

Since Y_{i+1} is independent from $(Y_1, \ldots, Y_i)$, repeated application of (3.19) shows that for $n \geq N$,

$$\mathcal{E}'\left[\sum_{i=2}^{n-1} D_i^n\right] < (n-2)\left(\frac{2K_2}{n^{1+\gamma/\alpha}} + \frac{\epsilon}{n}\right) < \frac{2K_2}{n^{\gamma/\alpha}} + \epsilon \tag{3.20}$$

and

$$\mathcal{E}'\left[\sum_{i=2}^{n-1} D_i^n\right] > -(n-2)\left(\frac{2K_2}{n^{1+\gamma/\alpha}} + \frac{\epsilon}{n}\right) > -\frac{2K_2}{n^{\gamma/\alpha}} - \epsilon. \tag{3.21}$$

We only need to combine our bounds above and invoke hypothesis (ii) to finish the proof. By (3.15), (3.16), (3.18), (3.20), and (3.21),

$$\begin{aligned}
&\mathcal{E}'\left[v\left(1, B_n S_n\right)\right] - v\left(0,0\right) \\
&= \mathcal{E}'\left[v\left(1, B_n S_n\right) - v\left(\frac{n-1}{n}, B_n S_n\right) + \sum_{i=2}^{n-1} C_i^n + \sum_{i=2}^{n-1} D_i^n + v\left(\frac{1}{n}, B_n S_2\right) - v\left(0,0\right)\right] \\
&\leq \mathcal{E}'\left[\left|v\left(1, B_n S_n\right) - v\left(\frac{n-1}{n}, B_n S_n\right)\right|\right] + \sum_{i=2}^{n-1} \mathcal{E}'\left[\left|C_i^n\right|\right] + \mathcal{E}'\left[\sum_{i=2}^{n-1} D_i^n\right] \\
&\qquad + \mathcal{E}'\left[\left|v\left(\frac{1}{n}, B_n S_2\right) - v\left(0,0\right)\right|\right] \\
&< \left(K_1\sqrt{\frac{1}{n}}\right) + \left(K_2\left[\left(\frac{1}{n}\right)^{\gamma/\alpha} + B_n^\gamma \mathcal{E}'\left[\left|Y_1\right|^\gamma\right]\right]\right) + \left(\frac{2K_2}{n^{\gamma/\alpha}} + \epsilon\right) \\
&\qquad + \left(K_1\sqrt{\frac{1}{n}} + 2\mathrm{Lip}\left(\psi\right) B_n \mathcal{E}'\left[\left|Y_1\right|\right]\right)
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{E}'\left[v\left(1, B_n S_n\right)\right] - v\left(0,0\right) \\
&\quad > -\left(K_1\sqrt{\frac{1}{n}}\right) - \left(K_2\left[\left(\frac{1}{n}\right)^{\gamma/\alpha} + B_n^\gamma \mathcal{E}'\left[\left|Y_1\right|^\gamma\right]\right]\right) - \left(\frac{2K_2}{n^{\gamma/\alpha}} + \epsilon\right) \\
&\qquad - \left(K_1\sqrt{\frac{1}{n}} + 2\mathrm{Lip}\left(\psi\right) B_n \mathcal{E}'\left[\left|Y_1\right|\right]\right)
\end{aligned}$$

for $n \geq N$. Since $\epsilon > 0$ is arbitrary and $\lim_{n\to\infty} B_n = 0$, we have

$$\lim_{n\to\infty} \mathcal{E}'\left[v\left(1, B_n S_n\right)\right] = v\left(0,0\right).$$

$\square$

3.4 Examples

Example III.13. X_1 is in its own domain of normal attraction. While this follows directly from the α-stability of X_1, we will demonstrate this using Theorem III.11 as well in order to unpack our main result.

Let $\psi \in C_{b.Lip}(\mathbb{R})$ and u be defined by

$$u(t,x) = \mathcal{E}\left[\psi\left(x + X_t\right)\right]$$

on $[0,\infty) \times \mathbb{R}$. If $\tilde{X}_1$ is an independent copy of X_1, then

$$\begin{aligned}
\mathcal{E}\left[\psi\left(aX_1 + b\tilde{X}_1\right)\right] &= \mathcal{E}\left[\mathcal{E}\left[\psi\left(ax + \left(b^\alpha\right)^{\frac{1}{\alpha}}\tilde{X}_1\right)\right]_{x=X_1}\right] \\
&= \mathcal{E}\left[u\left(b^\alpha, aX_1\right)\right] \\
&= u\left(a^\alpha + b^\alpha, 0\right) \\
&= \mathcal{E}\left[\psi\left(\left(a^\alpha + b^\alpha\right)^{\frac{1}{\alpha}} X_1\right)\right]
\end{aligned}$$

for any a, $b \geq 0$ by Lemmas III.6 and III.7, i.e., X_1 is α-stable. Exercise 3.20 in [194] implies that the same relation actually holds for a broader class of maps. In particular,

$$\begin{aligned}
2^{\frac{1}{\alpha}}\mathcal{E}\left[X_1\right] &= \mathcal{E}\left[\mathcal{E}\left[x + \tilde{X}_1\right]_{x=X_1}\right] \\
&= \mathcal{E}\left[X_1 + \mathcal{E}\left[X_1\right]\right] \\
&= 2\mathcal{E}\left[X_1\right],
\end{aligned}$$

so

$$\mathcal{E}\left[X_1\right] = 0.$$

It follows similarly that

$$\mathcal{E}\left[-X_1\right] = 0.$$

We know

$$\mathcal{E}\left[|X_1|\right] < \infty$$

from Lemma III.8.

To check the final hypothesis, let $0 < h < 1$ and v be the unique viscosity solution of (3.3). Then for all $(t, x) \in [0, 1] \times \mathbb{R}$,

$$\begin{aligned}
& n\left|\mathcal{E}\left[\delta_{B_n X_1} v\left(t, x\right)\right] - \left(\frac{1}{n}\right) \sup_{k_\pm \in K_\pm} \left\{\int_{\mathbb{R}} \delta_z v\left(t, x\right) F_{k_\pm}\left(dz\right)\right\}\right| \\
&= n\left|\mathcal{E}\left[v\left(t, x + B_n X_1\right)\right] - v\left(t, x\right) + \left(\frac{1}{n}\right) \partial_t v\left(t, x\right)\right| \\
&= n\left|v\left(t - \frac{1}{n}, x\right) - v\left(t, x\right) + \left(\frac{1}{n}\right) \partial_t v\left(t, x\right)\right| \\
&= \left|\frac{v\left(t - \frac{1}{n}, x\right) - v\left(t, x\right)}{1/n} + \partial_t v\left(t, x\right)\right| \\
&\leq \frac{K_2}{n^{\gamma/\alpha}}
\end{aligned}$$

by (3.12), (3.17), and Lemma III.7. Here $b = 1$ or, equivalently,

$$B_n = \frac{1}{n^{1/\alpha}}.$$

Abusing notation, Theorem III.11 shows that

$$B_n S_n \xrightarrow{d} X_1$$

as $n \to \infty$.

Example III.14. Up to some "uniformity" assumptions, this example has a straightforward interpretation.

> Let the uncertainty subset of distributions (see [194]) of Y_1 be given by $\{\mathbb{P}_\theta : \theta \in \Theta\}$. If for all $\theta \in \Theta$, a classical random variable with distribution $\mathbb{P}_\theta$ is in the domain of normal attraction of a classical α-stable random variable with triplet θ, then Y_1 is in the domain of normal attraction of X_1.

Let b, $M > 0$ and f be a nonnegative function on $\mathbb{N}$ tending to zero as $n \to \infty$. For each $k_\pm \in K_\pm$, let $W_{k_\pm}$ be a classical random variable such that

(i) $W_{k_\pm}$ has mean zero;

(ii) $W_{k_\pm}$ has a cdf $F_{W_{k_\pm}}$ of the form

$$F_{W_{k_\pm}}(z) = \begin{cases} \left[b^\alpha (k_-/\alpha) + \beta_{1,k_\pm}(z)\right] \frac{1}{|z|^\alpha} & z < 0 \\ 1 - \left[b^\alpha (k_+/\alpha) + \beta_{2,k_\pm}(z)\right] \frac{1}{z^\alpha} & z > 0 \end{cases} \tag{3.22}$$

for some continuously differentiable functions $\beta_{1,k_\pm}$ on $(-\infty, 0]$ and $\beta_{2,k_\pm}$ on $[0, \infty)$ with

$$\lim_{z\to-\infty} \beta_{1,k_\pm}(z) = \lim_{z\to\infty} \beta_{2,k_\pm}(z) = 0;$$

(iii) the following quantities are all less than M:

$$\left|\int_{-\infty}^{-1} \frac{\beta_{1,k_\pm}(z)}{(-z)^\alpha} dz\right| \quad , \quad \left|\int_{-1}^{0} \frac{\beta'_{1,k_\pm}(z)}{(-z)^{\alpha-1}} dz\right| \quad , \quad \int_{-1}^{0} \frac{\left|-\beta'_{1,k_\pm}(z) z + \alpha\beta_{1,k_\pm}(z)\right|}{(-z)^{\alpha-1}} dz,$$

$$\left|\int_{1}^{\infty} \frac{\beta_{2,k_\pm}(z)}{z^\alpha} dz\right| \quad , \quad \left|\int_{0}^{1} \frac{\beta'_{2,k_\pm}(z)}{z^{\alpha-1}} dz\right| \quad , \quad \int_{0}^{1} \frac{\left|-\beta'_{2,k_\pm}(z) z + \alpha\beta_{2,k_\pm}(z)\right|}{z^{\alpha-1}} dz; \text{ and}$$

(iv) the following quantities are less than $f(n)$ for all n:

$$\left|\beta_{2,k_\pm}\left(B_n^{-1}\right)\right| \quad , \quad \int_{1}^{\infty} \frac{\left|\beta_{2,k_\pm}(B_n^{-1} z)\right|}{z^\alpha} dz \quad , \quad \int_{0}^{1} \frac{\left|\beta_{2,k_\pm}(B_n^{-1} z)\right|}{z^{\alpha-1}} dz$$

$$\left|\beta_{1,k_\pm}\left(-B_n^{-1}\right)\right| \quad , \quad \int_{-\infty}^{-1} \frac{\left|\beta_{1,k_\pm}(B_n^{-1} z)\right|}{(-z)^\alpha} dz \quad , \quad \int_{-1}^{0} \frac{\left|\beta_{1,k_\pm}(B_n^{-1} z)\right|}{(-z)^{\alpha-1}} dz.$$

Note that by (ii) alone, the terms in (iii) are finite and the terms in (iv) approach zero as $n \to \infty$. In other words, the content of (iii) and (iv) is that uniform bounds and minimum rates of convergence exist.

Define an operator $\mathcal{E}'$ on a space $\mathcal{H}$ of suitable functions by

$$\mathcal{E}'[\varphi] = \sup_{k_\pm \in K_\pm} \int_{\mathbb{R}} \varphi(z)\, dF_{W_{k_\pm}}(z)$$

for all $\varphi \in \mathcal{H}$. The exact composition of $\mathcal{H}$ is irrelevant for our purposes here. Clearly, $(\mathbb{R}, \mathcal{H}, \mathcal{E}')$ is a sublinear expectation space.

Let Y_1 be the random variable on this space defined by

$$Y_1(x) = x$$

for all $x \in \mathbb{R}$. We will use Theorem III.11 to show that

$$B_n S_n \xrightarrow{d} X_1$$

as $n \to \infty$. Most of the difficulties have already been addressed during our discussion of the classical case in Section 3.3.

Since each $W_{k_\pm}$ has mean zero,

$$\mathcal{E}'[Y_1] = \sup_{k_\pm \in K_\pm} \int_{\mathbb{R}} z\, dF_{W_{k_\pm}}(z) = 0$$

and

$$\mathcal{E}'[-Y_1] = \sup_{k_\pm \in K_\pm} \int_{\mathbb{R}} -z\, dF_{W_{k_\pm}}(z) = 0.$$

After recalling that $K_\pm \subset (\lambda, \Lambda)$, (iii) gives

$$\mathcal{E}'[|Y_1|] < \infty$$

using (3.4) and (3.22). Observe that we are solving (3.22) for the obvious expressions to obtain uniform bounds on the terms

$$\left|\beta_{2,k_\pm}(1)\right| \quad , \quad \left|\beta_{1,k_\pm}(-1)\right| \quad , \quad \left|\int_0^1 \frac{b^\alpha k_+ + \alpha \beta_{2,k_\pm}(z)}{z^\alpha}\, dz\right|$$

and

$$\left|\int_{-1}^0 \frac{b^\alpha k_- + \alpha \beta_{1,k_\pm}(z)}{(-z)^\alpha}\, dz\right|.$$

To check the remaining hypothesis, let $0 < h < 1$, $\psi \in C_{b.Lip}(\mathbb{R})$, and v be the unique viscosity solution of (3.3). The techniques of (3.5) demonstrate that

$$\begin{aligned} &n \left| \mathcal{E}' \left[\delta_{B_n Y_1} v(t,x) \right] - \left(\frac{1}{n} \right) \sup_{k_\pm \in K_\pm} \left\{ \int_{\mathbb{R}} \delta_z v(t,x) F_{k_\pm}(dz) \right\} \right| \\ &\leq \left(\frac{1}{b^\alpha} \right) \sup_{k_\pm \in K_\pm} \left| \int_{\mathbb{R}} \delta_z v(t,x) \left(\frac{\beta'_{1,k_\pm}(B_n^{-1} z) \left| B_n^{-1} z \right| + \alpha \beta_{1,k_\pm}(B_n^{-1} z)}{|z|^{\alpha+1}} \mathbf{1}_{(-\infty,0)}(z) \right. \right. \\ &\qquad \left. \left. + \frac{-\beta'_{2,k_\pm}(B_n^{-1} z) \left| B_n^{-1} z \right| + \alpha \beta_{2,k_\pm}(B_n^{-1} z)}{|z|^{\alpha+1}} \mathbf{1}_{(0,\infty)}(z) \right) dz \right| \end{aligned}$$

for $(t,x) \in [0,1] \times \mathbb{R}$ and $n \geq 1$. Combining (3.6), (3.7), (3.8), and (3.9) with (iii) and (iv) proves that this last expression approaches zero in the required way.

3.5 Appendix: Interior Regularity Theory Background

Interior regularity theory for fully nonlinear integro-differential equations is rich and well-developed. Before describing the results that we need for our proof, we provide a short discussion of the literature. Readers new to this field are encouraged to first consult [12] for an introduction.

Some results and methods from the interior regularity theory for PDEs can be imported to the nonlocal case after minor modifications. For other aspects of the theory, this is false. As described in Section 2 of [217], a Hölder estimate and the Harnack inequality appear together in the local setting; however, there are nonlocal equations for which a Hölder estimate holds in the absence of the Harnack inequality. A partial list of other ways that nonlocal results can significantly differ from their local counterparts can be found in [12].

Early work on the regularity of integro-differential equations focused on equations in divergence form. A survey of these results is contained in [146]. For equations in nondivergence form, [39] contains the first Harnack inequality and Hölder estimate.

The equations studied in [39] are of the form

$$\int_{\mathbb{R}^d} \left[w\left(x+z\right)-w\left(x\right)-z\nabla w\left(x\right)\mathbf{1}_{B_1}\left(z\right)\right]k\left(x,z\right)\,dz=0,$$

where k is a kernel such that

$$k\left(x,z\right)=k\left(x,-z\right) \tag{3.23}$$

and

$$\frac{\lambda_1}{\left|z\right|^{d+\alpha_1}}\leq k\left(x,z\right)\leq\frac{\Lambda_1}{\left|z\right|^{d+\alpha_1}} \tag{3.24}$$

for some constants λ_1, $\Lambda_1>0$ and $\alpha_1\in(0,2)$. For a review of the extensions of this initial work, see [146].

The Hölder estimate in [39] blows up as $\alpha_1\to 2$. Many other early estimates share this feature. The first paper to prove a Hölder estimate and Harnack inequality without this property is [65]. The equations are of the form

$$\inf_r\sup_s\left\{\int_{\mathbb{R}^d}\left[w\left(x+z\right)-w\left(x\right)-z\nabla w\left(x\right)\mathbf{1}_{B_1}\left(z\right)\right]k^{rs}\left(z\right)\,dz\right\}=0 \tag{3.25}$$

for kernels k^{rs} depending only on z and satisfying (3.23), (3.24), and an additional smoothness condition. More precisely, for some fixed positive constants ρ and C,

$$\int_{\mathbb{R}^d\setminus B_\rho}\frac{\left|k\left(z\right)-k\left(z-\epsilon\right)\right|}{\left|\epsilon\right|}\,dz\leq C$$

whenever

$$\left|\epsilon\right|<\frac{\rho}{2}.$$

The paper culminates in a $C^{1,\gamma}$ estimate for the solution of (3.25).

These findings have been extended in a number of ways. For instance, references such as [221], [218], [155], [157], and [156] study equations with nonsymmetric kernels, i.e., kernels that do not satisfy (3.23). Other examples of recent work include [66], [219], and [153].

We now collect the definitions and results from [157] and [156] that we need for our proof. These references describe properties of the solutions to a broad class of nonlocal fully nonlinear parabolic equations of the form

$$\partial_t w\left(t,x\right) - Iw\left(t,x\right) = f\left(t\right).$$

Due to the general nature of these equations, [157] and [156] are quite technical. Since (3.1) is an easy case of the equations studied in these papers, we will simplify this material and present only the version that we need for our argument.

Notation III.15. Let

$$\mathfrak{C}_{\tau,r}\left(t,x\right) := \left(t-\tau,t\right] \times \left(x-r,x+r\right).$$

We write $\mathfrak{C}_{\tau,r}$ for the cylinder $\mathfrak{C}_{\tau,r}\left(0,0\right)$. For suitable functions w, let

$$\begin{aligned}
&\tilde{\delta}_z w\left(t,x\right) := w\left(t,x+z\right) - w\left(t,x\right) - \partial_x w\left(t,x\right)\mathbf{1}_{(-1,1)}\left(z\right) z;\\
&\|w\|_{L^1(\nu)} := \int_{\mathbb{R}} \left|w\left(z\right)\right| \min\left(1,|z|^{-1-\alpha}\right)\, dz;\ \text{and}\\
&[w]_{C^{0,1}((t_0,t_1]\mapsto L^1(\nu))} := \sup_{(t-\tau,t]\subseteq(t_0,t_1]} \frac{\left\|w\left(t,\cdot\right) - w\left(t-\tau,\cdot\right)\right\|_{L^1(\nu)}}{\tau}.
\end{aligned}$$

We also let

$$b_{k_\pm} := \left(k_- - k_+\right)\int_1^\infty \frac{dz}{z^\alpha}$$

for all $k_\pm \in K_\pm$.

In the literature, one also works frequently with cylinders of the form

$$\left(t-\tau^\alpha,t\right] \times \left(x-r,x+r\right)$$

due to their convenient scaling properties. We introduce

$$\|\cdot\|_{L^1(\nu)}$$

and

$$[\,\cdot\,]_{C^{0,1}((t_0,t_1]\mapsto L^1(\nu))}$$

due to their role in upcoming Hölder estimates, namely, Lemmas III.18 and III.19. The symbols $\tilde{\delta}_z$ and $b_{k_\pm}$ facilitate the identification of (3.1) with the equations studied [157] and [156]. Observe that for all $k_\pm \in K_\pm$ and suitable functions w,

$$\int_{\mathbb{R}} \delta_z w\,(t,x)\, F_{k_\pm}\,(dz) = b_{k_\pm}\,\partial_x w\,(t,x) + \int_{\mathbb{R}} \tilde{\delta}_z w\,(t,x)\, F_{k_\pm}\,(dz)\,. \tag{3.26}$$

Definition III.16. Since $K_\pm \subset (\lambda, \Lambda)$, we can pick $\beta > 0$ such that

$$\sup_{k_\pm \in K_\pm} \left\{ \sup_{r\in(0,1)} \left\{ r^{\alpha-1} \left| b_{k_\pm} + \int_{(-1,1)\setminus(-r,r)} z F_{k_\pm}\,(dz) \right| \right\} \right\} \le \beta.$$

Let $\mathcal{L}_0$ be the family of operators

$$w\,(t,x) \mapsto b\,\partial_x w\,(t,x) + \int_{\mathbb{R}} \tilde{\delta}_z w\,(t,x)\, \frac{k\,(z)}{|z|^{1+\alpha}}\, dz,$$

where k is a kernel and b is a constant such that $\lambda \le k \le \Lambda$ and

$$\sup_{r\in(0,1)} r^{\alpha-1} \left| b + \int_{(-1,1)\setminus(-r,r)} \frac{z k\,(z)}{|z|^{1+\alpha}}\, dz \right| \le \beta.$$

We say that an operator in $\mathcal{L}_0$ is in $\mathcal{L}_1$ if

$$|\partial_z k\,(z)| \le \frac{\Lambda}{|z|},$$

and an operator in $\mathcal{L}_1$ is in $\mathcal{L}_2$ if

$$\left|\partial^2_{zz} k\,(z)\right| \le \frac{\Lambda}{|z|^2}.$$

The stronger regularity requirements on the kernels (in $\mathcal{L}_2$, say, compared to those in $\mathcal{L}_0$) give rise to stronger regularity results. All of the operators

$$w\,(t,x) \mapsto b_{k_\pm}\,\partial_x w\,(t,x) + \int_{\mathbb{R}} \tilde{\delta}_z w\,(t,x)\, F_{k_\pm}\,(dz)$$

are in each of these families. As we will soon see in (3.27), we will be especially interested in the operator I defined by

$$Iw\left(t,x\right)=\inf_{k_{\pm}\in K_{\pm}}\left\{b_{k_{\pm}}\,\partial_{x}w\left(t,x\right)+\int_{\mathbb{R}}\tilde{\delta}_{z}w\left(t,x\right)F_{k_{\pm}}\left(dz\right)\right\}.$$

I is a specific case of an *extremal operator*.

Definition III.17. For a collection of operators $\mathcal{L}\subseteq\mathcal{L}_0$, define the *extremal operators* $\mathcal{M}_{\mathcal{L}}^{+}$ and $\mathcal{M}_{\mathcal{L}}^{-}$ by

$$\mathcal{M}_{\mathcal{L}}^{+}=\sup_{L\in\mathcal{L}}L\qquad\text{and}\qquad\mathcal{M}_{\mathcal{L}}^{-}=\inf_{L\in\mathcal{L}}L.$$

I has a number of other key properties including the following.[4]

(i) $I0=0$.

(ii) I is *uniformly elliptic* with respect to $\mathcal{L}_j$, i.e.,

$$\mathcal{M}_{\mathcal{L}_j}^{-}\left(w_1-w_2\right)\leq Iw_1-Iw_2\leq\mathcal{M}_{\mathcal{L}_j}^{+}\left(w_1-w_2\right).$$

(iii) I is *translation invariant*, i.e.,

$$I\left(w\left(t_0+\cdot,x_0+\cdot\right)\right)\left(t,x\right)=\left(Iw\right)\left(t_0+t,x_0+x\right).$$

(i) is trivial. See Section 2 of [156] for (ii). Since I has constant coefficients, we get (iii). We highlight these classes of operators and properties of I for the convenience of the reader comparing the next three results to their original versions (see Theorem 2.3 in [156] for Lemma III.18; Theorems 1.1, 2.4, and 2.5 in [156] for Lemma III.19; and Theorem 3.3 in [157] for Lemma III.20).[5]

[4]Though we will not emphasize this point, we remark in passing that $Iw\left(t,x\right)$ is well-defined for any $w\left(t,\cdot\right)\in C^{1,1}\left(x\right)\cap L^{1}\left(\nu\right)$ (see Section 2 of [156]).

[5]A number of related results exist in the literature. We mention only a small sample. Theorem 12.1 in [65], Theorem 1.1 in [217], and Theorem 7.1 in [218] are C^{γ} estimates along the lines of Lemma III.18. Theorem 8.1 in [218], Theorem 13.1 in [65], Theorem 1.1 in [66], and Theorem 1.1 in [219] contain $C^{1,\gamma}$ or $C^{\alpha+\gamma}$ estimates similar to those in Lemma III.19. Like Lemma III.20, Theorem 5.9 in [65] and Lemma 3.2 in [220] investigate the difference of viscosity solutions.

Lemma III.18. *Let w satisfy*

$$\partial_t w - M^+_{\mathcal{L}_0} w \leq 0$$

$$\partial_t w - M^-_{\mathcal{L}_0} w \geq 0$$

in the viscosity sense on $\mathfrak{C}_{1,1}$. There is some $\gamma \in (0,1)$ and $C > 0$ depending only on λ, Λ, and β such that for every (t_0, x_0), $(t_1, x_1) \in \mathfrak{C}_{1/2,1/2}$,

$$\frac{|w(t_0, x_0) - w(t_1, x_1)|}{\left(|t_0 - t_1|^{1/\alpha} + |x_0 - x_1|\right)^\gamma} \leq C \|w\|_{L^1((-1,0] \mapsto L^1(\nu))}.$$

Lemma III.19. *Let w satisfy*

$$\partial_t w - Iw = 0$$

in the viscosity sense on $\mathfrak{C}_{1,1}$. There is some $\gamma \in (0,1)$ and $C > 0$ depending only on λ, Λ, and β such that for every (t_0, x_0), $(t_1, x_1) \in \mathfrak{C}_{1/2,1/2}$,

$$|\partial_x w(t_0, x_0)| + \frac{|\partial_x w(t_0, x_0) - \partial_x w(t_1, x_1)|}{\left(|t_0 - t_1|^{1/\alpha} + |x_0 - x_1|\right)^\gamma} \leq C \|w\|_{L^1((-1,0] \mapsto L^1(\nu))}$$

and

$$|\partial_t w(t_0, x_0)| + \frac{|\partial_t w(t_0, x_0) - \partial_t w(t_1, x_1)|}{\left(|t_0 - t_1|^{1/\alpha} + |x_0 - x_1|\right)^\gamma} \leq C [w]_{C^{0,1}((-1,0] \mapsto L^1(\nu))}.$$

We also have

$$\|w\|_{C^{\alpha+\gamma}\left(\mathfrak{C}_{1/2,1/2}\right)} \leq C \left(\|w\|_{L^1((-1,0] \mapsto L^1(\nu))} + \left[w \mathbf{1}_{(-1,1)^c}\right]_{C^{0,1}((-1,0] \mapsto L^1(\nu))}\right).$$

Lemma III.20. *Let w_1, w_2 satisfy*

$$\partial_t w_i - I w_i = 0$$

in the viscosity sense on some domain Ω. Then

$$\partial_t (w_1 - w_2) - M^+_{\mathcal{L}_0} (w_1 - w_2) \leq 0$$

$$\partial_t (w_1 - w_2) - M^-_{\mathcal{L}_0} (w_1 - w_2) \geq 0$$

also holds in the viscosity sense on Ω.

We will need one more result (for the original version, see Lemma 5.6 and the proof of Corollary 5.7 in [64]). It is the key to a standard technique from the literature allowing one to repeatedly apply an estimate such as Lemma III.18 in order to obtain a higher regularity estimate.

Lemma III.21. *Let* $0 < \beta_1 \leq 1$, $0 < \beta_2 < 1$, $L > 0$, *and* $w \in L^\infty([-1,1])$ *satisfy*

$$\|w\|_{L^\infty([-1,1])} \leq L.$$

For $0 < |h_0| \leq 1$, *define* w_{β_1,h_0} *by*

$$w_{\beta_1,h_0}(x) = \frac{w(x+h_0) - w(x)}{|h_0|^{\beta_1}}$$

for all $x \in I_{h_0}$, *where* $I_{h_0} = [-1, 1-h_0]$ *if* $h_0 > 0$ *and* $I_{h_0} = [-1-h_0, 1]$ *if* $h_0 < 0$. *Suppose that*

$$w_{\beta_1,h_0} \in C^{\beta_2}(I_{h_0})$$

and

$$\|w_{\beta_1,h_0}\|_{C^{\beta_2}(I_{h_0})} \leq L$$

for any $0 < |h_0| \leq 1$.

(i) If $\beta_1 + \beta_2 < 1$, *then*

$$w \in C^{\beta_1+\beta_2}([-1,1])$$

and

$$\|w\|_{C^{\beta_1+\beta_2}([-1,1])} \leq CL.$$

(ii) If $\beta_1 + \beta_2 > 1$ *and* $\beta_1 \neq 1$, *then*

$$w \in C^{0,1}([-1,1])$$

and

$$\|w\|_{C^{0,1}([-1,1])} \leq CL.$$

(iii) If $\beta_1 = 1$*, then* $w \in C^{1,\beta_2}([-1,1])$ *and*

$$\|w\|_{C^{1,\beta_2}([-1,1])} \leq CL.$$

In any of these cases, C *depends only on* $\beta_1 + \beta_2$*.*

We will often apply these results on different domains than we have listed above without comment. For instance, we might use Lemma III.19 on $\mathfrak{C}_{1,1}(t,x)$ or Lemma III.21 on an arbitrary closed interval. These "new" results are obtained merely by translating or rescaling, both standard routines in the literature. As an example of such an operation, notice that if w satisfies

$$\partial_t w - Iw = 0$$

in the viscosity sense on $(t_1, t_2] \times \Omega$, then $\tilde{w}$ defined by

$$\tilde{w}(t,x) = w(r^\alpha t + t_0, rx + x_0)$$

satisfies

$$\partial_t \tilde{w} - I\tilde{w} = 0$$

in the viscosity sense on

$$\left(\frac{t_1 - t_0}{r^\alpha}, \frac{t_2 - t_0}{r^\alpha}\right] \times \frac{\Omega - x_0}{r}$$

(see Section 2.1.1 of [157]). Further information can be found in [157], [156], and [64].

3.6 Appendix: Proof of Proposition III.10

In the hope of keeping the number of constants in our argument at a reasonable level, we will not issue a new subscript each time we introduce a new constant B

below. Also, we will write $\bar{u}$ instead of $-u$. From (3.1) and (3.26), $\bar{u}$ is a viscosity solution of

$$\partial_t \bar{u}(t,x) - I\bar{u}(t,x) = 0, \quad (t,x) \in (0,\infty) \times \mathbb{R}$$
$$\bar{u}(0,x) = -\psi(x), \quad x \in \mathbb{R}. \tag{3.27}$$

It suffices to show that parts (i)-(iv) of Proposition III.10 hold for $\bar{u}$ and (3.27).

The quantities

$$[\bar{u}]_{C^{0,1}((t_0,t_1]\mapsto L^1(\nu))}$$

play a crucial role in Lemma III.19, so our first goal will be to control them for t_0 greater than some positive number. We will do this by showing that $\bar{u}$ is uniformly Lipschitz as a function of time for times above some lower bound. Achieving a Lipschitz estimate can be done using a standard strategy. Specifically, we will begin by obtaining an initial $C^{\gamma/\alpha}$ estimate from Lemma III.18. Lemma III.20 will allow us to apply Lemma III.18 to get a $C^{\gamma/\alpha}$ estimate for the incremental quotients of $\bar{u}$. Then Lemma III.21 will give that $\bar{u}$ is $C^{2\gamma/\alpha}$ in time. We will repeat these steps to show that $\bar{u}$ is $C^{3\gamma/\alpha}$ in time, $C^{4\gamma/\alpha}$ in time, and so on until we conclude that $\bar{u}$ is Lipschitz in time.

Since

$$\mathcal{M}^-_{\mathcal{L}_0} w \leq Iw \leq \mathcal{M}^+_{\mathcal{L}_0} w,$$

$\bar{u}$ satisfies

$$\partial_t \bar{u} - M^+_{\mathcal{L}_0} \bar{u} \leq 0$$
$$\partial_t \bar{u} - M^-_{\mathcal{L}_0} \bar{u} \geq 0$$

in the viscosity sense on $(0,\infty)\times\mathbb{R}$. For any $\bar{t}>1$,

$$\begin{aligned}\|\bar{u}\left(\bar{t}+\cdot,\cdot\right)\|_{L^1((-1,0]\mapsto L^1(\nu))} &= \int_{-1}^{0}\int_{\mathbb{R}}\left|\bar{u}\left(\bar{t}+t,z\right)\right|\min\left(1,|z|^{-1-\alpha}\right)\,dz\,dt\\ &\leq \|\psi\|_{L^\infty(\mathbb{R})}\int_{-1}^{0}\int_{\mathbb{R}}\min\left(1,|z|^{-1-\alpha}\right)\,dz\,dt\end{aligned}$$

by Lemma III.9. Lemma III.18 implies that for some B, $\gamma>0$,

$$\frac{\left|\bar{u}\left(t_0,x_0\right)-\bar{u}\left(t_1,x_1\right)\right|}{\left(\left|t_0-t_1\right|^{1/\alpha}+\left|x_0-x_1\right|\right)^{\gamma}}\leq B. \tag{3.28}$$

for every (t_0,x_0), $(t_1,x_1)\in\mathfrak{C}_{1/2,1/2}\left(\bar{t},\bar{x}\right)$ with $\bar{t}>1$.

For $0<|h_0|<1/2$, define $\bar{u}_{\gamma/\alpha,h_0}$ by

$$\bar{u}_{\gamma/\alpha,h_0}\left(t,x\right)=\frac{\bar{u}\left(t+h_0,x\right)-\bar{u}\left(t,x\right)}{|h_0|^{\gamma/\alpha}}$$

for all $(t,x)\in[1/2,\infty)\times\mathbb{R}$. Then

$$\left\|\bar{u}_{\gamma/\alpha,h_0}\right\|_{L^\infty((1,\infty)\times\mathbb{R})}\leq B$$

by (3.28). Hence,

$$\left\|\bar{u}_{\gamma/\alpha,h_0}\left(\bar{t}+\cdot,\cdot\right)\right\|_{L^1((-1,0]\mapsto L^1(\nu))}\leq B\int_{\mathbb{R}}\min\left(1,|z|^{-1-\alpha}\right)\,dz$$

for any $\bar{t}>2$.

Notice that

$$\partial_t\bar{u}\left(\cdot+h_0,\cdot\right)-I\bar{u}\left(\cdot+h_0,\cdot\right)=0$$

in the viscosity sense on $(1/2,\infty)\times\mathbb{R}$ because (3.27) has constant coefficients. Lemma III.20 implies that

$$\partial_t\bar{u}_{\gamma/\alpha,h_0}-M^+_{\mathcal{L}_0}\bar{u}_{\gamma/\alpha,h_0}\leq 0$$

$$\partial_t\bar{u}_{\gamma/\alpha,h_0}-M^-_{\mathcal{L}_0}\bar{u}_{\gamma/\alpha,h_0}\geq 0$$

in the viscosity sense on $(1/2,\infty)\times\mathbb{R}$. For some B,

$$\frac{\left|\bar{u}_{\gamma/\alpha,h_0}(t_0,x_0)-\bar{u}_{\gamma/\alpha,h_0}(t_1,x_1)\right|}{\left(|t_0-t_1|^{1/\alpha}+|x_0-x_1|\right)^{\gamma}}\leq B$$

for every (t_0,x_0), $(t_1,x_1)\in\mathfrak{C}_{1/2,1/2}(\bar{t},\bar{x})$ with $\bar{t}>2$ by Lemma III.18.

Lemma III.21 shows that for a small r_1 (less than $1/4$), we can find B such that

$$\bar{u}(\cdot,\bar{x})\in C^{2\gamma/\alpha}([\bar{t}-r_1,\bar{t}+r_1])$$

and

$$\|\bar{u}(\cdot,\bar{x})\|_{C^{2\gamma/\alpha}([\bar{t}-r_1,\bar{t}+r_1])}\leq B \tag{3.29}$$

for $\bar{t}>2$.

Due to Lemma III.21, assume without loss of generality that α/γ is not an integer. Starting from the incremental quotient

$$\frac{\bar{u}(t+h_0,x)-\bar{u}(t,x)}{|h_0|^{2\gamma/\alpha}},$$

we can use these steps to produce a $C^{3\gamma/\alpha}$ estimate for $\bar{u}$ in time. By continuing to repeat this procedure, we will obtain a $C^{4\gamma/\alpha}$ estimate, a $C^{5\gamma/\alpha}$ estimate, and so on until we obtain a Lipschitz estimate for $\bar{u}$ in time. More precisely, we will find B and a small r_n such that

$$\bar{u}(\cdot,\bar{x})\in C^{0,1}([\bar{t}-r_n,\bar{t}+r_n])$$

and

$$\|\bar{u}(\cdot,\bar{x})\|_{C^{0,1}([\bar{t}-r_n,\bar{t}+r_n])}\leq B$$

for $\bar{t}>\lceil\alpha/\gamma\rceil$.

For t_0, $t_1>\lceil\alpha/\gamma\rceil$,

$$\begin{aligned}|\bar{u}(t_0,x_0)-\bar{u}(t_1,x_0)| &\leq |\bar{u}(s_0,x_0)-\bar{u}(s_1,x_0)|+\cdots+|\bar{u}(s_{N-1},x_0)-\bar{u}(s_N,x_0)|\\ &\leq B|s_0-s_1|+\cdots+B|s_{N-1}-s_N|\\ &= B|t_0-t_1|\end{aligned}$$

where $t_0 = s_0$, $t_1 = s_N$, and $s_i < s_{i+1} \le s_i + 2r_n$. This indicates that

$$\bar{u}\left(\cdot,\bar{x}\right) \in C^{0,1}\left(\left(\lceil\alpha/\gamma\rceil,\infty\right)\right)$$

and

$$\left\|\bar{u}\left(\cdot,\bar{x}\right)\right\|_{C^{0,1}\left(\left(\lceil\alpha/\gamma\rceil,\infty\right)\right)} \le B.$$

Then t_0, $t_1 > \lceil\alpha/\gamma\rceil$ implies

$$\begin{aligned}
\left[\bar{u}\mathbf{1}_{(-1,1)^c}\right]_{C^{0,1}((t_0,t_1]\mapsto L^1(\nu))} &\le \left[\bar{u}\right]_{C^{0,1}((t_0,t_1]\mapsto L^1(\nu))} \\
&= \sup_{(t-\tau,t]\subseteq(t_0,t_1]} \frac{\left\|\bar{u}\left(t,\cdot\right)-\bar{u}\left(t-\tau,\cdot\right)\right\|_{L^1(\nu)}}{\tau} \\
&\le B\int_{\mathbb{R}} \min\left(1,|z|^{-1-\alpha}\right)\, dz.
\end{aligned}$$

Lemma III.19 gives that for $\bar{t} > \lceil\alpha/\gamma\rceil$,

$$\left|\partial_x\bar{u}\left(t_0,x_0\right)\right| + \frac{\left|\partial_x\bar{u}\left(t_0,x_0\right)-\partial_x\bar{u}\left(t_1,x_1\right)\right|}{\left(\left|t_0-t_1\right|^{1/\alpha}+\left|x_0-x_1\right|\right)^{\gamma}} \le B \tag{3.30}$$

and

$$\left|\partial_t\bar{u}\left(t_0,x_0\right)\right| + \frac{\left|\partial_t\bar{u}\left(t_0,x_0\right)-\partial_t\bar{u}\left(t_1,x_1\right)\right|}{\left(\left|t_0-t_1\right|^{1/\alpha}+\left|x_0-x_1\right|\right)^{\gamma}} \le B \tag{3.31}$$

for every (t_0,x_0), $(t_1,x_1) \in \mathfrak{C}_{1/2,1/2}\left(\bar{t},\bar{x}\right)$. It also shows that

$$\left\|\bar{u}\right\|_{C^{\alpha+\gamma}\left(\mathfrak{C}_{1/2,1/2}(\bar{t},\bar{x})\right)} \le B. \tag{3.32}$$

After suitably rescaling, we see that these inequalities actually hold for $\bar{t} > (1+h)/2$. Part (i) of Proposition III.10 then follows from (3.30) and (3.31), while part (iii) follows from (3.32). From (3.31) and a simple covering argument, we know that as long as the distance between x_0 and x_1 is under some arbitrary bound, we can find B such that

$$\left|\partial_t\bar{u}\left(t,x_0\right)-\partial_t\bar{u}\left(t,x_1\right)\right| \le B\left|x_0-x_1\right|^{\gamma}$$

for $t \in [h, h+1]$. Since $\partial_t \bar{u}$ is bounded on $[h, h+1] \times \mathbb{R}$, we can drop the distance constraint and get the second inequality in part (ii). A similar covering argument finishes the proof of the first inequality and yields part (ii) of Proposition III.10.

It remains to prove part (iv). In this case, the equation for $\bar{u}$ is

$$\partial_t \bar{u}(t,x) - b_{k_\pm}\, \partial_x \bar{u}(t,x) - \int_{\mathbb{R}} \tilde{\delta}_z \bar{u}(t,x)\, F_{k_\pm}(dz) = 0, \quad (t,x) \in (0,\infty) \times \mathbb{R}$$

$$\bar{u}(0,x) = -\psi(x), \quad x \in \mathbb{R}. \tag{3.33}$$

Since $\bar{u}$ is a classical solution of this equation on $[h,\infty) \times \mathbb{R}$, $\bar{u}(\cdot, \bar{x} + \cdot)$ also classically satisfies

$$\partial_t \bar{u}(\cdot, \bar{x}+\cdot) - b_{k_\pm}\, \partial_x \bar{u}(\cdot, \bar{x}+\cdot) - \int_{\mathbb{R}} \tilde{\delta}_z \bar{u}(\cdot, \bar{x}+\cdot)\, F_{k_\pm}(dz) = 0$$

on $[h,\infty) \times \mathbb{R}$. Then

$$\hat{u}_{h_0}(t,x) := \frac{\bar{u}(t, x+h_0) - \bar{u}(t,x)}{|h_0|}$$

is a classical solution of (3.33) on $[h,\infty) \times \mathbb{R}$ as well.

Lemma III.9 implies that

$$\|\hat{u}_{h_0}(\bar{t}+\cdot, \cdot)\|_{L^1((-1,0] \mapsto L^1(\nu))} \leq \operatorname{Lip}(\psi) \int_{\mathbb{R}} \min\left(1, |z|^{-1-\alpha}\right) dz$$

for $\bar{t} > 1$. By Lemma III.19, it follows that for some B,

$$|\partial_x \hat{u}_{h_0}(t_0,x_0)| + \frac{|\partial_x \hat{u}_{h_0}(t_0,x_0) - \partial_x \hat{u}_{h_0}(t_1,x_1)|}{\left(|t_0 - t_1|^{1/\alpha} + |x_0 - x_1|\right)^{\gamma}} \leq B$$

for every (t_0, x_0), $(t_1, x_1) \in \mathfrak{C}_{1/2,1/2}(\bar{t}, \bar{x})$ with $\bar{t} > 1$. Rewriting this in terms of $\bar{u}$, we see that we have found a γ-Hölder estimate for

$$\frac{\partial_x \bar{u}(t_0, x+h_0) - \partial_x \bar{u}(t_0, x)}{|h_0|}.$$

By Lemma III.21, $\partial^2_{xx} \bar{u}$ exists and is bounded on $(1/2, \infty) \times \mathbb{R}$. By rescaling, we get that this actually holds on $[h, h+1] \times \mathbb{R}$. $\square$

CHAPTER IV

Index Tracking Near Rebalance Dates

4.1 Introduction

An index fund is distinguished by its objective. In the words of [230], "the goal of an index fund is to track the performance of a specific market benchmark as closely as possible". One might think that the difference between the fund's return and the benchmark's return, also known as the *tracking difference*, should be the key metric here. The most important metric in practice is actually the *tracking error*, i.e., the standard deviation of the difference between the fund's return and the benchmark's return.

Despite its significance, this measurement possesses a few curious features. A purely theoretical remark is that a fund which maintains a constant difference between its return and the benchmark's return has a tracking error of zero, regardless of the value of the difference. Of more practical significance is the observation that tracking error penalizes outperformance. In fact, "index funds do not attempt to outperform their benchmark" according to [105].

Regardless, index funds are enormously popular: [236] estimates that nearly 20% of what is invested in stock funds lies in an index fund. Endorsements come from luminaries including Warren Buffett. In [63], he writes that "most investors, both institutional and individual, will find that the best way to own common stocks is through an index fund that charges minimal fees". Empirical studies support this strategy too, for example, [104] finds that index funds beat comparable actively managed porfolios over 80% of the time. Their tax efficiency, broadly diversified portfolios, low fees, and clear investment objectives are some of their most frequently cited advantages.

Managers of index funds still face significant difficulties, one of the biggest of which is that indexes occasionally rebalance or reconstitute. During a reconstitution, the securities in an index can be reweighted, current securities can be removed (*index deletions*), or new securities can be added (*index additions*). Rebalancing procedures among indexes vary widely, both in the specifics of the procedures and even their transparency to the public.

The Russell Indexes use the especially well-understood construction methods detailed in [14]. For instance, the Russell 3000 Index includes the top 3,000 U.S. stocks based on market capitalization. To complete this year's reconstitution, Russell Investments will begin an initial ranking of all U.S. stocks after the close on May 29, 2015. A preliminary list of index deletions and additions will be released to the public on June 12, 2015. Updated lists will be released on the next two Fridays, and the reconstitution will take effect after the close on June 26, 2015.

The strategies employed by many index fund managers are also quite well-understood. In [148], Kim and Oikonomou write that

> "with a primary objective of replicating the performance of any given benchmark, managers have a much higher incentive to minimize tracking error than to take greater risks to improve returns. [...] Passive index funds with the tightest tracking error allowances typically choose to wait

> until the last moment on the effective day of change before trading, leading to inflated prices on the purchases of those stocks being added to the index (and substantially lower prices on those to be removed)".

The combination of widely known index rebalancing methods and index tracking strategies presents ripe opportunities for predatory traders. Some funds such as the Aviva Investors Index Opportunities Fund are very open about their intentions in this regard and are openly designed to profit from the behavior of index funds near rebalance dates. It is determined in [83] that investors in funds tracking the S&P 500 and the Russell 2000 lose between one and two billion dollars per year to predatory trading of this kind. Perhaps even more startling is the finding in [67] that a buy-and-hold portfolio outperforms the annually rebalanced index by an average of 17.29% over five years.

The goal of this paper is to construct a model for the problem faced by an index tracker during a reconstitution. We want to take into account varying tracking error constraints, market characteristics, and predatory trading activities. We hope to understand how these variables affect the returns and optimal strategies for the index tracker and a potential predator, as well as the prices of the index securities. To the best of our knowledge, ours is the first predatory trading model to explicitly incorporate tracking error considerations.

In our model, the index tracker's tracking error constraint winds up being an extremely strong condition. As this bound becomes tighter, the index tracker's strategy approaches an instantaneous liquidation at the terminal time, regardless of the market conditions. Consequently, the price of an index deletion during the reconstitution is significantly deflated. This appears to reflect practical experience.

As we will see, the index tracker in our model sometimes benefits from the presence of the predator. The intuition behind this observation is that the negative effects due to the "predator's role as a predator" are outweighed by the positive effects due to the "predator's role as a liquidity provider", as long as the index tracker is eventually selling rapidly enough. This is never the case in the related one-period models of [69] and [216].

We give a short overview of the literature in Section 4.2. Our model is described in Section 4.3. In Section 4.4, we describe the mathematical details underlying our numerical work. Section 4.5 contains our numerical examples.

4.2 Previous Work

A number of empirical studies have been completed on various aspects of index rebalancing. Some of the key papers include [164], [82], [179], [84], [166], [180], [67], [83], [119], and [112]. Another helpful reference is the undergraduate thesis [240]. In Section 4.1, we have discussed the findings in these papers that we are most interested in reproducing with our model.

Many theoretical studies of predatory trading in general have also been done. Early models are proposed in [62] and [26]. In [69] and [216], risk-neutral agents maximize expected revenues, and open-loop Nash equilibria are determined using an Almgren-Chriss market impact model. This framework is extended to the case of multiple assets in [86]. Schied and Zhang study an open-loop Nash equilibrium where the players conduct either a CARA utility maximization or a certain mean-variance optimization in [213]. Carmona and Yang numerically analyze a closed-loop Nash equilibrium with one distressed trader and one predator, where the predator can trade over a longer time horizon in [72]. In [215], Schöneborn derives both open-loop and closed-loop strategies for two players using a discrete-time limit order book model extending that of [178]. Unlike previous studies, [170] investigates a game with asymmetric information.

4.3 Our Model

We work on some finite time horizon, say $[0, T]$. One index tracker and some noise traders are active in the market throughout this period. Another strategic trader, who we will view as a predator, is also sometimes present. It will occasionally be convenient to refer to the index tracker and the predator as Player 1 and Player 2, respectively. The description of our model will be less awkward if we act as if the predator is always active. The case where the index tracker and the noise traders are alone in the market can be obtained by merely dropping all of the quantities related to the predator.

Our market contains only two assets: one risky asset and one risk-free asset (cash). We assume that the risk-free interest rate is zero. Of course, an index consists of many securities in practice, but we will assume for simplicity that our index consists of only a single asset. More precisely, at $t = 0$, only the risky asset is a member of the index. At $t = T$, the following reconstitution takes effect: the risky asset is deleted and the risk-free asset is added.

We assume that the index is guaranteed to rebalance in this way and that both the index tracker and the predator are aware of this at $t = 0$. This means that our time horizon should be viewed as being fairly short, say the last day of trading before the reconstitution. We remark that even with a transparent index such as the Russelll 2000, no such guarantee actually exists in practice. Also, while we could have replaced the risky asset with another risky asset to make our model more realistic, we opt against this for convenience. [1]

[1] Still, our preliminary (unpublished) investigations into the two risky asset case actually suggest that the results

Let $X_1(t)$ and $X_2(t)$ denote the risky asset holdings at time t of the index tracker and the predator, respectively. We require $X_1(t)$ and $X_2(t)$ to be continuously differentiable. The price at time t of the risky asset, $P(t)$, is given by the linear price impact model

$$P(t) = \tilde{P}(t) + \sum_{i=1}^{2} \gamma_i (X_i(t) - X_i(0)) + \sum_{i=1}^{2} \lambda_i \dot{X}_i(t) \tag{4.1}$$

for $t \in [0, T]$, where $\tilde{P}$ is an arithmetic Brownian motion without drift starting from $\bar{p}$. Here, γ_i is Player i's *permanent* price impact parameter, while λ_i is Player i's *temporary* price impact parameter.

This model for the asset price dynamics is identical to that used in many sources, e.g., [69] and [216], except that we allow the index tracker and the predator to have potentially different price impact parameters. One reason for this choice is that we can recover a simple version of the N-predator model in [69] and [216] by letting $\gamma_2 = N\gamma_1$ and $\lambda_2 = N\lambda_1$.

Index trackers do not necessarily own all of the securities in the index they are tracking. We strive to avoid such complications, so we require the index tracker to hold only the risky asset at $t = 0$ and only cash at $t = T$. We denote the index tracker's initial holdings in the risky asset by $\bar{x}$, so we have

$$X_1(0) = \bar{x} \tag{4.2}$$

and

$$X_1(T) = 0. \tag{4.3}$$

For the predator, we enforce the conditions

$$X_2(0) = 0 \tag{4.4}$$

are easily understood from our setup here anyway, at least when the assets are independent.

and

$$X_2(T) = 0. \tag{4.5}$$

The first condition seems justified by the observation that the predator only trades in the risky asset to take advantage of the index tracker's distressed circumstances. This suggests that the predator should not have any initial holdings in the asset. Because many index reconstitutions take effect at the market close on the effective date, it is reasonable to view T as a market closing time. The second condition then reflects the fact that predatory traders often eschew having open positions at the end of the trading day.

At $t = 0$, both the index tracker and the predator will choose their deterministic trading rates over the entire time horizon. We denote these functions by $\dot{X}_1$ and $\dot{X}_2$, respectively. The expected revenue for Player i becomes

$$-\int_0^T \dot{X}_i(t)\left[\bar{p} + \gamma_1\left(X_1(t) - \bar{x}\right) + \gamma_2 X_2(t) + \lambda_1 \dot{X}_1(t) + \lambda_2 \dot{X}_2(t)\right] dt.$$

Observe that this is deterministic as well.

In many predatory trading models, e.g., [69] and [216], the traders maximize their expected revenues. We will assume that the predator has this objective throughout; however, this might not be an appropriate assumption for the index tracker. Let us examine this claim more carefully.

In our setup, the index tracker's tracking error on $[0, T]$ is

$$\sqrt{\operatorname{Var}\left(\frac{P(T) - \bar{p}}{\bar{p}} + \frac{\int_0^T \dot{X}_1(t)\, P(t)\, dt + \bar{x}\bar{p}}{\bar{x}\bar{p}}\right)}. \tag{4.6}$$

As an aside, to avoid an artificial jump in the benchmark's value, we consider the index's holdings in the risk-free asset to be worth $P(T)$ after time T. A similar

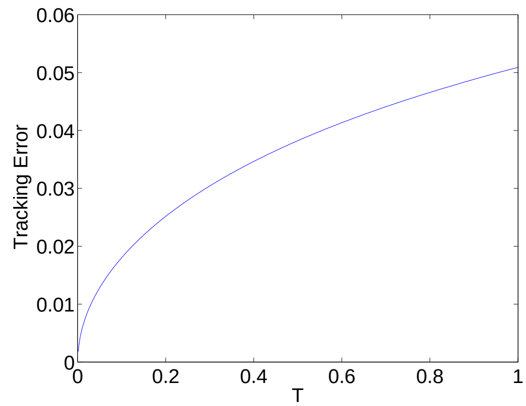

Figure 4.1: Depiction of the index tracker's tracking error (see Section 4.3).

normalization occurs in practice for the same reason. The issue is that (4.6) may not be small enough if the index tracker is maximizing expected revenue. In the plastic market case from [216], i.e., letting

$$\bar{x} = 1, \quad \bar{p} = 10, \quad T = 1,$$

$$\gamma_1 = \gamma_2 = 3, \quad \text{and} \quad \lambda_1 = \lambda_2 = 1,$$

we have

$$\sqrt{\operatorname{Var}\left(\frac{P(T) - \bar{p}}{\bar{p}} + \frac{\int_0^T \dot{X}_1(t)\, P(t)\, dt + \bar{x}\bar{p}}{\bar{x}\bar{p}}\right)} = 0.05.$$

Tracking error allowances can be much smaller.

An obvious remedy is to shorten the time horizon. This appears to reflect how index trackers deal with this issue in practice. Figure 4.1 shows the index tracker's tracking error as a function of T using the rest of the parameter values just mentioned. Unfortunately, the index tracker's new strategy is only guaranteed to be optimal on the shortened time horizon, not the original time horizon. Also, arbitrarily shortening the index tracker's time horizon should not necessarily shorten the predator's time horizon as well.

One could allow the predator to trade in a period before the index tracker enters the market. This is the approach taken in [240]. The index tracker's strategy is still not guaranteed to be optimal on the original time horizon, so we feel an alternative setup is needed.

A second remedy might be to assume that the index tracker's objective is to minimize (4.6). After attempting to carry out the program in Section 4.4, we found that the only possible solution features

$$X_1(t) = \bar{x} H(T - t),$$

where H is the Heaviside step function. Intuitively, this solution means that the index tracker instantaneously liquidates the risky asset at the terminal time. While this solution also appears to reflect the behavior of index trackers in practice, it is problematic in our setup. The main concern is that $P(T)$ is no longer defined: the index tracker's trading rate is

$$\dot{X}_1(t) = -\bar{x}\delta(T - t),$$

after all. Another objection is that the index tracker's risky asset position is not continuously differentiable.

In an attempt to avoid the occurrence of delta functions, a third remedy might be to assume that the index tracker's objective is to minimize (4.6) and that all trading rates are valued in a given compact set, say $[-M, M]$. After attempting to apply the techniques of Section 4.4, we found that the only possible solution is as follows: the index tracker initially does not trade at all and then sells the risky asset at the maximum allowable speed. Intuitively, the index tracker uses the strategy which most closely approximates the strategy from the previous setup, which is perhaps not a surprise. The primary concern about this approach is that many important quantities in the problem now depend on a fairly arbitrary trading speed bound. A secondary issue is that the index tracker's risky asset position is not continuously differentiable.

We have now arrived at the remedy which we use in this paper. For some $\alpha > 0$, the index tracker's objective is to maximize

$$-\int_0^T \dot{X}_1(t)\left[\bar{p} + \gamma_1(X_1(t) - \bar{x}) + \gamma_2 X_2(t) + \lambda_1 \dot{X}_1(t) + \lambda_2 \dot{X}_2(t)\right] dt$$

such that

$$\sqrt{\mathrm{Var}\left(\frac{P(T) - \bar{p}}{\bar{p}} + \frac{\int_0^T \dot{X}_1(t) P(t)\, dt + \bar{x}\bar{p}}{\bar{x}\bar{p}}\right)} = \alpha. \tag{4.7}$$

The interpretation here is that the index tracker is maximizing expected revenue while maintaining a specific tracking error. Based on Section 4.5, we suspect that if we replace "=" in (4.7) with "≤", we will get an equivalent problem.

A nice feature of this setup is that the solutions in the one-period models of [69] and [216] are recovered as special cases. Of course, the value of α which allows us to recover their result depends on the values of all of the other parameters.

4.4 Mathematical Details

As it stands, our game is formulated in a rather unusual way due to (4.7). This problem is easily solved. We introduce the extra state variables X_3 and X_4 satisfying

$$\begin{aligned} \dot{X}_3(t) &= t\dot{X}_1(t) \\ \dot{X}_4(t) &= X_3(t)\,\dot{X}_1(t) \end{aligned} \tag{4.8}$$

with initial data

$$\begin{aligned} X_3(0) &= 0 \\ X_4(0) &= 0. \end{aligned} \tag{4.9}$$

These variables are defined precisely so that we can rewrite the tracking error in terms of the terminal values of the state variables. Specifically, a quick calculation shows that the index tracker's tracking error constraint is now given by

$$\sqrt{\frac{T}{\bar{p}^2} + \frac{2X_3\left(T\right)}{\bar{x}\bar{p}^2} + \frac{2X_4\left(T\right)}{\bar{x}^2\bar{p}^2}} = \alpha. \tag{4.10}$$

4.4.1 No Predator Case

We begin by analyzing the case when the predator is absent from the market. Note that we merely drop the quantities related to the predator throughout this discussion and retain all of the other notation. We will use standard necessary optimality conditions based upon the maximum principle, e.g., see [31], to produce a candidate strategy for the index tracker.

Introduce the costate function

$$p\left(t\right) = \left(p_1\left(t\right), p_3\left(t\right), p_4\left(t\right)\right)$$

for $t \in [0, T]$ and the constant multipliers ν_1, ν_2. The Hamiltonian for the index tracker is given by

$$\begin{aligned} H\left(X\left(t\right), \dot{X}_1\left(t\right), p\left(t\right), t\right) = \dot{X}_1\left(t\right) \Big[& \bar{p} + \gamma_1\left(X_1\left(t\right) - \bar{x}\right) + \lambda_1 \dot{X}_1\left(t\right) \\ & + p_1\left(t\right) + t p_3\left(t\right) + X_3\left(t\right) p_4\left(t\right) \Big] \end{aligned}$$

on $[0, T]$. The costate function satisfies the equations

$$\begin{aligned} \dot{p}_1\left(t\right) &= -\gamma_1 \dot{X}_1\left(t\right) \\ \dot{p}_3\left(t\right) &= -p_4\left(t\right) \dot{X}_1\left(t\right) \\ \dot{p}_4\left(t\right) &= 0 \end{aligned}$$

with terminal data

$$
\begin{aligned}
p_1(T) &= \nu_1 \\
p_3(T) &= \frac{2\nu_2}{\bar{x}\bar{p}^2} \\
p_4(T) &= \frac{2\nu_2}{\bar{x}^2\bar{p}^2}.
\end{aligned}
$$

This means that

$$
\begin{aligned}
p_1(t) &= -\gamma_1 X_1(t) + \nu_1 \\
p_3(t) &= -\frac{2\nu_2}{\bar{x}^2\bar{p}^2}\left(X_1(t) - \bar{x}\right) \\
p_4(t) &= \frac{2\nu_2}{\bar{x}^2\bar{p}^2}
\end{aligned}
$$

for $t \in [0, T]$.

By the stationarity condition,

$$
\begin{aligned}
0 &= \bar{p} + \gamma_1\left(X_1(t) - \bar{x}\right) + 2\lambda_1\dot{X}_1(t) + p_1(t) + tp_3(t) + X_3(t)\,p_4(t) \\
&= \bar{p} - \gamma_1\bar{x} + \nu_1 + 2\lambda_1\dot{X}_1(t) - \frac{2\nu_2}{\bar{x}^2\bar{p}^2}\int_0^t [t-s]\,\dot{X}_1(s)\,ds.
\end{aligned}
$$

Hence,

$$
\dot{X}_1(t) = -\frac{\bar{p} - \gamma_1\bar{x} + \nu_1}{2\lambda_1} + \frac{\nu_2}{\lambda_1\bar{x}^2\bar{p}^2}\int_0^t [t-s]\,\dot{X}_1(s)\,ds.
$$

This is a non-homogeneous Volterra integral equation of the second kind. From Section 2.1 of [202],

$$
\begin{aligned}
\dot{X}_1(t) &= \frac{-\bar{p} + \gamma_1\bar{x} - \nu_1}{2\lambda_1} + \sqrt{\frac{\nu_2}{\lambda_1\bar{x}^2\bar{p}^2}}\left(\frac{-\bar{p} + \gamma_1\bar{x} - \nu_1}{2\lambda_1}\right)\int_0^t \sinh\left[\sqrt{\frac{\nu_2}{\lambda_1\bar{x}^2\bar{p}^2}}\,(t-s)\right]ds \\
&= \left(\frac{-\bar{p} + \gamma_1\bar{x} - \nu_1}{2\lambda_1}\right)\cosh\left(t\sqrt{\frac{\nu_2}{\lambda_1\bar{x}^2\bar{p}^2}}\right) \qquad (4.11)
\end{aligned}
$$

on $[0, T]$. Note that if $\nu_2 < 0$, we could equivalently write

$$
\dot{X}_1(t) = \left(\frac{-\bar{p} + \gamma_1\bar{x} - \nu_1}{2\lambda_1}\right)\cos\left(t\sqrt{-\frac{\nu_2}{\lambda_1\bar{x}^2\bar{p}^2}}\right).
$$

Let

$$A = \frac{-\bar{p} + \gamma_1 \bar{x} - \nu_1}{2\lambda_1}$$

and

$$B = \sqrt{\frac{\nu_2}{\lambda_1 \bar{x}^2 \bar{p}^2}}.$$

If $B = 0$, then integrating (4.11) gives

$$X_1(t) = \bar{x} + At$$

on $[0, T]$. By (4.3),

$$A = -\frac{\bar{x}}{T}.$$

Continuing, we get

$$\begin{aligned} \dot{X}_1(t) &= -\frac{\bar{x}}{T} \\ X_1(t) &= \bar{x}\left[1 - \frac{t}{T}\right] \\ X_3(t) &= -\frac{\bar{x}t^2}{2T} \\ X_4(t) &= \frac{\bar{x}^2 t^3}{6T^2}, \end{aligned} \tag{4.12}$$

which leads to

$$\alpha^2 = \frac{T}{3\bar{p}^2}. \tag{4.13}$$

Now assume that $B \neq 0$. Integrating (4.11) gives

$$X_1(t) = \bar{x} + \frac{A \sinh(Bt)}{B}$$

on $[0, T]$. By (4.3),

$$A = -\frac{\bar{x}B}{\sinh(BT)}.$$

After simplifying, (4.8) implies that

$$\begin{aligned}
\dot{X}_1(t) &= -\frac{\bar{x}B\cosh(Bt)}{\sinh(BT)} \\
X_1(t) &= \bar{x}\left[1-\frac{\sinh(Bt)}{\sinh(BT)}\right] \\
X_3(t) &= -\frac{\bar{x}\left[Bt\sinh(Bt)-\cosh(Bt)+1\right]}{B\sinh(BT)} \\
X_4(t) &= \frac{\bar{x}^2\left[2Bt\cosh^2(Bt)-3\sinh(Bt)\cosh(Bt)+4\sinh(Bt)-3Bt\right]}{2B\left[\cosh(2BT)-1\right]}
\end{aligned} \tag{4.14}$$

By substituting our new expressions into (4.10) and simplifying again, we arrive at

$$\alpha^2 = -\frac{4BT\exp(2BT)-\exp(4BT)+1}{2B\bar{p}^2\left[\exp(2BT)-1\right]^2}. \tag{4.15}$$

From (4.12), (4.14), and (4.15), we see that $\dot{X}_1$ does not depend on the price impact parameters γ_1 and λ_1.

Notice that (4.12) would also be the optimal strategy if the index tracker maximized expected revenue and did not face a tracking error constraint, e.g., see page 2244 of [69]. The solution in that case depends only on $\bar{x}$ and T. While our solution also depends on $\bar{x}$ and T, in general, it depends on $\bar{p}$ (and α) as well.

The intuition here seems to be roughly as follows. As $\bar{p}$ increases, the effect of the index tracker's choices on both her own return and the return of the risky asset should decrease. Assuming the index tracker has a tracking error constraint, this should make the constraint easier to satisfy and give the index tracker greater flexbility when selecting a strategy. In particular, the index tracker should take $\bar{p}$ into account. On the other hand, if the index tracker is only trying to maximize her expected revenue, she should be primarily worried about how much her trading causes the price of the risky asset to drop in an absolute sense. Because this is independent of the risky asset's initial price (in our model, at least), the index tracker's strategy should be independent of $\bar{p}$ in this case.

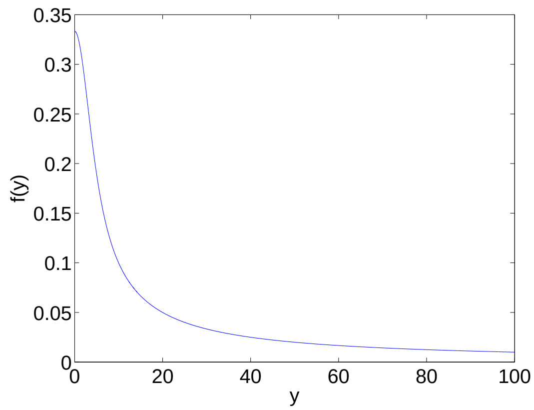

Figure 4.2: Depiction of $f(\cdot)$ (see Subsection 4.4.1).

Whether the index tracker maximizes expected revenue (only) or maximizes expected revenue with a tracking error constraint, the values of the price impact parameters γ_1 and λ_1 do not affect the optimal solution when the predator is absent from the market.

To the best of our knowledge, (4.15) cannot be solved explicitly for B. Since we believe it reveals all of the salient features of the problem, we will only informally analyze this equation. By changing variables, note that solving (4.15) is equivalent to solving

$$\frac{\alpha^2 \bar{p}^2}{T} = -\frac{2y \exp(y) - \exp(2y) + 1}{y \left[\exp(y) - 1\right]^2}. \tag{4.16}$$

for y. Denote by f the map

$$y \mapsto -\frac{2y \exp(y) - \exp(2y) + 1}{y \left[\exp(y) - 1\right]^2}.$$

Based on our work above, the only relevant arguments for f lie along either the positive real axis or the positive imaginary axis.

Figure 4.2 is a graph of $f(y)$ for positive real y, while Figure 4.3 is a graph of $f(iy)$ for positive real y. We can draw the following conclusions:

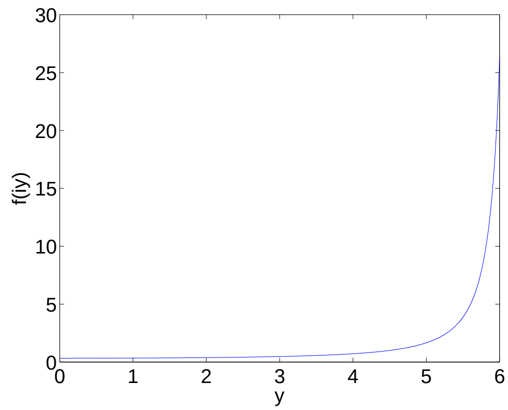

Figure 4.3: Depiction of $f(i\cdot)$ (see Subsection 4.4.1).

(i)

$$\lim_{y\downarrow 0} f(y) = \lim_{y\downarrow 0} f(iy) = \frac{1}{3},$$

(ii)

$$\lim_{y\uparrow +\infty} f(y) = 0,$$

(iii)

$$\lim_{y\uparrow 2\pi} f(iy) = +\infty,$$

(iv) $f(y)$ decreases as y increases for y on the positive real axis, and

(v) $f(iy)$ increases as y increases for $y \in (0, 2\pi)$.

Comparing these observations to (4.13) and (4.15), we see that

$$\begin{aligned}
\alpha^2 > \frac{T}{3\bar{p}^2} &\implies B \in i\mathbb{R}^+ \implies \nu_2 < 0\\
\alpha^2 = \frac{T}{3\bar{p}^2} &\implies B = 0 \implies \nu_2 = 0\\
\alpha^2 < \frac{T}{3\bar{p}^2} &\implies B \in \mathbb{R}^+ \implies \nu_2 > 0.
\end{aligned}$$

Of course, the case where

$$\alpha^2 > \frac{T}{3\bar{p}^2}$$

is merely a curiosity. The index tracker would have a higher expected revenue and a lower tracking error by setting $B = 0$ instead, a far more preferable outcome in practice (this case corresponds to the strategy which maximizes expected revenue).

4.4.2 Predator Case

We now study the case where the predator is active in the market. We will use the same approach as in the previous subsection. The resulting expressions and equations are far more complicated here due to the appearance of the predator. For this reason, our discussion will be limited.

For $j = 1, 2$, we introduce the costate function

$$p_j(t) = (p_{j,1}(t), p_{j,2}(t), p_{j,3}(t), p_{j,4}(t))$$

on $[0, T]$. We also let ν_1, ν_2, and ν_3 be constant multipliers. The Hamiltonian for Player j is defined by

$$\begin{aligned} &H_j\left(X(t), \dot{X}_1(t), \dot{X}_2(t), p_j(t), t\right) \\ &\quad = \dot{X}_j(t)\left[\bar{p} + \gamma_1(X_1(t) - \bar{x}) + \gamma_2 X_2(t) + \lambda_1 \dot{X}_1(t) + \lambda_2 \dot{X}_2(t)\right] \\ &\qquad + \dot{X}_1(t)\left[p_{j,1}(t) + t p_{j,3}(t) + X_3(t) p_{j,4}(t)\right] + \dot{X}_2(t) p_{j,2}(t) \end{aligned}$$

for $j = 1, 2$ and $t \in [0, T]$. The costate function j satisfies

$$\begin{aligned} \dot{p}_{j,1}(t) &= -\gamma_1 u_j \\ \dot{p}_{j,2}(t) &= -\gamma_2 u_j \\ \dot{p}_{j,3}(t) &= -\dot{X}_1 p_{j,4} \\ \dot{p}_{j,4}(t) &= 0 \end{aligned}$$

subject to the terminal conditions

$$\begin{aligned}
p_{1,1}(T) &= \nu_1 \\
p_{1,2}(T) &= 0 \\
p_{1,3}(T) &= \frac{2\nu_2}{\bar{x}\bar{p}^2} \\
p_{1,4}(T) &= \frac{2\nu_2}{\bar{x}^2\bar{p}^2}
\end{aligned}$$

and

$$\begin{aligned}
p_{2,1}(T) &= 0 \\
p_{2,2}(T) &= \nu_3 \\
p_{2,3}(T) &= 0 \\
p_{2,4}(T) &= 0.
\end{aligned}$$

In particular,

$$\begin{aligned}
p_{1,1}(t) &= -\gamma_1 \int_0^t \dot{X}_1(s)\ ds - \gamma_1\bar{x} + \nu_1 \\
p_{1,3}(t) &= -\frac{2\nu_2}{\bar{x}^2\bar{p}^2} \int_0^t \dot{X}_1(s)\ ds \\
p_{1,4}(t) &= \frac{2\nu_2}{\bar{x}^2\bar{p}^2} \\
p_{2,2}(t) &= -\gamma_2 \int_0^t \dot{X}_2(s)\ ds + \nu_3
\end{aligned}$$

on $[0, T]$.

From the stationarity condition, we get

$$\begin{aligned}
0 &= \bar{p} + \gamma_1(X_1(t) - \bar{x}) + \gamma_2 X_2(t) + 2\lambda_1\dot{X}_1(t) + \lambda_2\dot{X}_2(t) + p_{1,1}(t) + tp_{1,3}(t) + X_3(t)\,p_{1,4}(t) \\
&= \bar{p} - \gamma_1\bar{x} + \nu_1 + \gamma_2 \int_0^t \dot{X}_2(s)\ ds + 2\lambda_1\dot{X}_1(t) + \lambda_2\dot{X}_2(t) - \frac{2\nu_2}{\bar{x}^2\bar{p}^2} \int_0^t [t-s]\,\dot{X}_1(s)\ ds
\end{aligned}$$

and

$$\begin{aligned} 0 &= \bar{p} + \gamma_1 \left(X_1 \left(t \right) - \bar{x} \right) + \gamma_2 X_2 \left(t \right) + \lambda_1 \dot{X}_1 \left(t \right) + 2 \lambda_2 \dot{X}_2 \left(t \right) + p_{2,2} \left(t \right) \\ &= \bar{p} + \gamma_1 \int_0^t \dot{X}_1 \left(s \right) \, ds + \lambda_1 \dot{X}_1 \left(t \right) + 2 \lambda_2 \dot{X}_2 \left(t \right) + \nu_3 \end{aligned} \tag{4.17}$$

for $t \in [0, T]$. This implies that

$$\begin{aligned} -\gamma_2 \int_0^t \dot{X}_2 \left(s \right) \, ds + \lambda_2 \dot{X}_2 \left(t \right) = -\nu_3 - \gamma_1 \bar{x} + \nu_1 + \lambda_1 \dot{X}_1 \left(t \right) - \gamma_1 \int_0^t \dot{X}_1 \left(s \right) \, ds \\ - \frac{2 \nu_2}{\bar{x}^2 \bar{p}^2} \int_0^t \left[t - s \right] \dot{X}_1 \left(s \right) \, ds. \end{aligned}$$

We can think of $\dot{X}_2$ as the solution to a non-homogeneous Volterra integral equation of the second kind involving $\dot{X}_1$. By Section 2.1 of [202], we have

$$\begin{aligned} \dot{X}_2 \left(t \right) = \frac{-\nu_3 - \gamma_1 \bar{x} + \nu_1}{\lambda_2} + \frac{\lambda_1}{\lambda_2} \dot{X}_1 \left(t \right) - \frac{\gamma_1}{\lambda_2} \int_0^t \dot{X}_1 \left(s \right) \, ds - \frac{2 \nu_2}{\lambda_2 \bar{x}^2 \bar{p}^2} \int_0^t \left[t - s \right] \dot{X}_1 \left(s \right) \, ds \\ + \frac{\gamma_2}{\lambda_2} \int_0^t \exp \left(\frac{\gamma_2 \left(t - r \right)}{\lambda_2} \right) \left[\frac{-\nu_3 - \gamma_1 \bar{x} + \nu_1}{\lambda_2} + \frac{\lambda_1}{\lambda_2} \dot{X}_1 \left(r \right) - \frac{\gamma_1}{\lambda_2} \int_0^r \dot{X}_1 \left(s \right) \, ds \right. \\ \left. - \frac{2 \nu_2}{\lambda_2 \bar{x}^2 \bar{p}^2} \int_0^r \left[r - s \right] \dot{X}_1 \left(s \right) \, ds \right] dr \end{aligned} \tag{4.18}$$

on $[0, T]$.

Define the auxiliary variable Y by

$$Y \left(t \right) = \int_0^t X_1 \left(s \right) \, ds$$

for $t \in [0,T]$. We can write (4.17) as

$$
\begin{aligned}
0 &= \bar{p} + \gamma_1 \int_0^t \dot{X}_1(s)\, ds + \lambda_1 \dot{X}_1(t) + \nu_3 \\
&\quad + 2\lambda_2 \Bigg(\frac{-\nu_3 - \gamma_1 \bar{x} + \nu_1}{\lambda_2} + \frac{\lambda_1}{\lambda_2}\dot{X}_1(t) - \frac{\gamma_1}{\lambda_2}\int_0^t \dot{X}_1(s)\, ds - \frac{2\nu_2}{\lambda_2 \bar{x}^2 \bar{p}^2}\int_0^t [t-s]\, \dot{X}_1(s)\, ds \\
&\qquad + \frac{\gamma_2}{\lambda_2}\int_0^t \exp\left(\frac{\gamma_2(t-r)}{\lambda_2}\right)\Bigg[\frac{-\nu_3 - \gamma_1\bar{x} + \nu_1}{\lambda_2} + \frac{\lambda_1}{\lambda_2}\dot{X}_1(r) - \frac{\gamma_1}{\lambda_2}\int_0^r \dot{X}_1(s)\, ds \\
&\qquad\qquad - \frac{2\nu_2}{\lambda_2 \bar{x}^2 \bar{p}^2}\int_0^r [r-s]\, \dot{X}_1(s)\, ds\Bigg] dr \Bigg) \\
&= \bar{p} - \gamma_1\left(\dot{Y}(t) - \bar{x}\right) + 3\lambda_1 \ddot{Y}(t) + \nu_3 + 2\left[-\nu_3 - \gamma_1\bar{x} + \nu_1\right] - \frac{4\nu_2}{\bar{x}^2\bar{p}^2}\left(-t\bar{x} + Y(t)\right) \\
&\quad + \frac{2\gamma_2}{\lambda_2}\Bigg(\frac{\lambda_2(\nu_1 - \nu_3)\left(\exp\left(\frac{\gamma_2 t}{\lambda_2}\right) - 1\right)}{\gamma_2} - \frac{2\lambda_2\nu_2\left(\lambda_2 - \lambda_2\exp\left(\frac{\gamma_2 t}{\lambda_2}\right) + \gamma_2 t\right)}{{\gamma_2}^2\,\bar{p}^2\bar{x}} \\
&\qquad + \int_0^t \exp\left(\frac{\gamma_2(t-r)}{\lambda_2}\right)\left[\lambda_1\ddot{Y}(r) - \gamma_1\dot{Y}(r) - \frac{2\nu_2}{\bar{x}^2\bar{p}^2}Y(r)\right] dr \Bigg)
\end{aligned}
$$

after integrating.

Let $\mathcal{L}$ denote the Laplace transform of Y. By taking Laplace transforms and solving for $\mathcal{L}$, we get

$$
\begin{aligned}
\mathcal{L}(s) = &\left[\frac{\bar{x}}{s^2\left(4\lambda_2\nu_2 + \gamma_2\gamma_1\bar{p}^2\bar{x}^2 + \gamma_2\lambda_1\bar{p}^2 s\bar{x}^2 + \gamma_1\lambda_2\bar{p}^2 s\bar{x}^2 - 3\lambda_2\lambda_1\bar{p}^2 s^2\bar{x}^2\right)}\right] \\
&\cdot \left[4\lambda_2\nu_2 - \gamma_2\bar{p}^3\bar{x} - \gamma_2\nu_3\bar{p}^2\bar{x} + \lambda_2\bar{p}^3 s\bar{x} + \gamma_2\gamma_1\bar{p}^2\bar{x}^2 + \gamma_2\lambda_1\bar{p}^2 s\bar{x}^2\right. \\
&\quad \left. - \gamma_1\lambda_2\bar{p}^2 s\bar{x}^2 - 3\lambda_2\lambda_1\bar{p}^2 s^2\bar{x}^2 + 2\lambda_2\nu_1\bar{p}^2 s\bar{x} - \lambda_2\nu_3\bar{p}^2 s\bar{x}\right].
\end{aligned}
$$

We will only state the next steps. The reason is that the expressions involved are enormous. While they can be easily obtained by following the procedure below using software such as MATLAB, they are so complicated that we believe it would not be beneficial for the reader to actually see them. We will see their important qualitative features during our discussion of the numerical examples soon.

Step 1: Find Y by taking the inverse Laplace transform of $\mathcal{L}$.

Step 2: Differentiate Y to obtain X_1.

Step 3: Differentiate X_1 to obtain $\dot{X}_1$.

Step 4: Use (4.8), (4.9), and the formula for $\dot{X}_1$ to obtain X_3 and X_4.

Step 5: Use (4.18) and the formula for $\dot{X}_1$ to obtain $\dot{X}_2$.

Step 6: Integrate $\dot{X}_2$ to obtain X_2.

All of the formulas above will be in terms of the constants ν_1, ν_2, and ν_3. It only remains to find these constants, which can be done numerically as follows.

Step 7: Use (4.3) and the formula for X_1 to obtain a formula for ν_1 in terms of ν_2 and ν_3.

Step 8: Use (4.5), the formula for X_2, and the formula for ν_1 from Step 7 to obtain a formula for ν_3 in terms of ν_2 alone.

Step 9: Use (4.10), the formulas for X_3 and X_4, and the formulas from Steps 7 and 8 to obtain an equation relating ν_2 to α, $\bar{p}$, $\bar{x}$, T, and the price impact parameters.

As in (4.15), it is not clear if the equation in Step 9 can be solved explicitly for ν_2; however, it is easy to do so numerically after selecting values for α, $\bar{p}$, $\bar{x}$, T, and the price impact parameters. We then use Steps 7 and 8 to get numerical values for ν_1 and ν_3. These values allow us to compute the remaining quantities of interest.

Before moving on to the numerical examples, we remark that $\dot{X}_1$ and $\dot{X}_2$ are of the form

$$\dot{X}_1(t) = \exp(C_1 t)\left[C_2 \sinh(C_3 t) + C_4 \cosh(C_3 t)\right]$$

and

$$\dot{X}_2(t) = \exp(C_1 t)[C_5 \sinh(C_3 t) + C_6 \cosh(C_3 t)] + C_7$$

for some constants $C_1, \ldots, C_7$.

4.5 Numerical Examples

For simplicity, we will first examine the case when the index tracker and the predator have the same price impact parameters. More precisely, we will use the parameters from the plastic market case in [216], i.e., we let

$$\bar{x} = 1, \quad \bar{p} = 10, \quad T = 1,$$

$$\gamma_1 = \gamma_2 = 3, \quad \text{and} \quad \lambda_1 = \lambda_2 = 1.$$

Recall that predatory trading has been observed in plastic markets (and in other types of markets as well). We will later vary the price impact parameters to understand their influence.

In all of our figures below, we will write "$N = 1$" or "$N = 0$" to indicate whether the predator is active or not, respectively. We only considered values of α that were below the value of α corresponding to the one-period model solution from [69] or [216], as these are the only solutions with potential relevance in practice. We also restricted ourselves to values of α for which the expected price process is positive on $[0, T]$, given our parameters.

4.5.1 Effects Due to the Tracking Error Constraint

Question IV.1. *What is the effect of the tracking error constraint α on the players' trading rates?*

Figure 4.4 shows the index tracker's trading rate in the case where the predator is active as a function of t for various values of α. The index tracker sells more

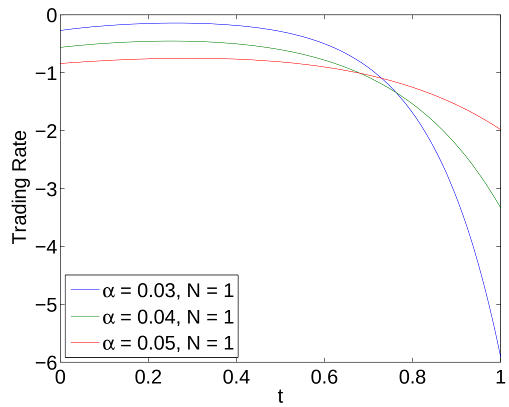

Figure 4.4: Depiction of the index tracker's trading rate (see Question IV.1).

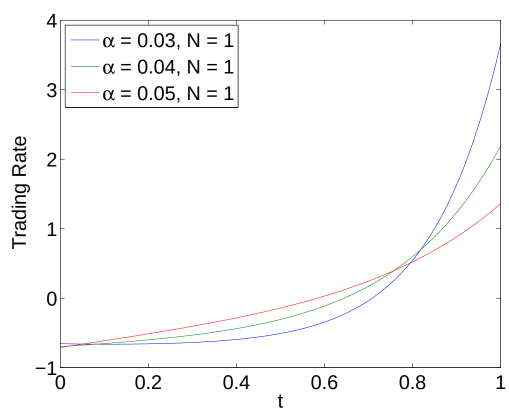

Figure 4.5: Depiction of the predator's trading rate (see Question IV.1).

rapidly for small times and more slowly for large times as α increases. By continuing to shrink α, the index tracker's trading rate appears to *converge* in an appropriate sense to

$$\dot{X}_1(t) = -\bar{x}\delta(T-t),$$

i.e., an instantaneous liquidation at the terminal time. This reflects practical experience, where tight tracking error constraints effectively force index fund managers to rebalance their portfolios at the close on the effective date of the reconstitution. Observe that the index tracker's trading rate retains the general features of the one-period model solution in [69] and [216], e.g., it is negative, concave, and has its maximum at an intermediate time.

In Figure 4.5, we have the predator's trading rate as a function of t for a few values of α. The predator sells more rapidly for (very) small times and buys more slowly for large times as α increases. The predator sells for a shorter period of time as α increases. Note that the predator's trading rate also retains important features of the one-period model solution in [69] and [216]. For instance, it is increasing, negative for small times, positive for large times, and convex. The predator's buying rate near the terminal time also rapidly increases when the index tracker's selling rate rapidly decreases near the terminal time, as in [69] and [216].

Question IV.2. *What is the effect of the tracking error constraint α on the expected price process?*

Figure 4.6 depicts the expected price process as a function of t in the case where the predator is active for several values of α. The expected price process decreases for small times but increases for large times as α increases. We also see a concavity changes as α increases. Specifically, as α decreases, it appears that the expected

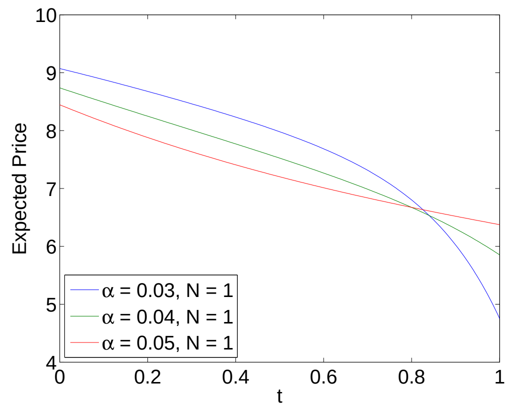

Figure 4.6: Depiction of the expected price (see Question IV.2).

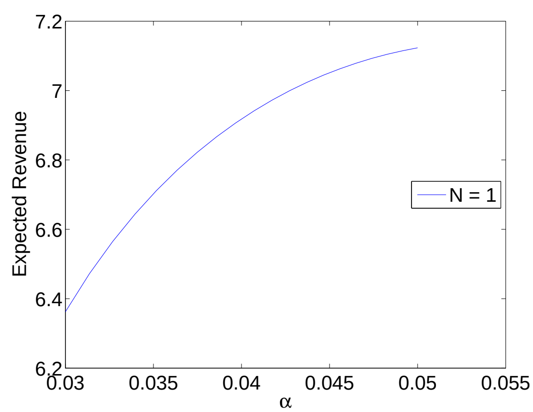

Figure 4.7: Depiction of the index tracker's expected revenue (see Question IV.3).

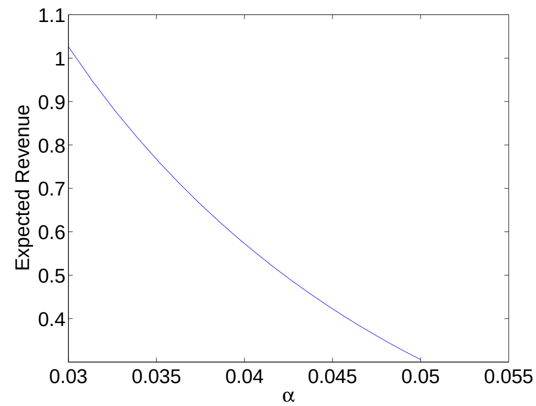

Figure 4.8: Depiction of the predator's expected revenue (see Question IV.3).

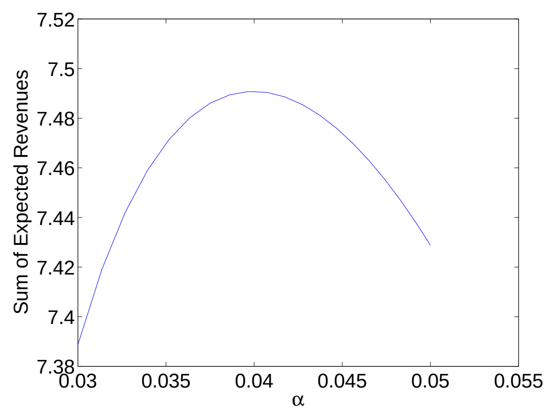

Figure 4.9: Depiction of the total expected revenue (see Question IV.3).

price process becomes concave. To the best of our knowledge, the expected price process in the one-period model of [69] and [216] is convex.

Question IV.3. *What is the effect of the tracking error constraint α on the players' expected revenues?*

Figure 4.7 shows the index tracker's expected revenue when the predator is active as a function of α. It appears to be increasing and concave.

In Figure 4.8, we have the predator's expected revenue as a function of α. It seems to be decreasing and convex.

Figure 4.9 depicts the sum of the index tracker's expected revenue (in the case where the predator is active) and the predator's expected revenue as a function of

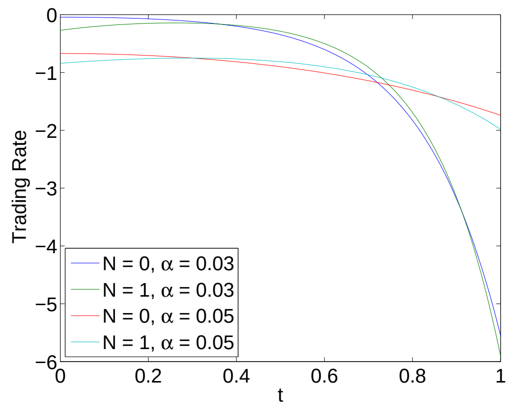

Figure 4.10: Depiction of the index tracker's trading rate (see Question IV.4).

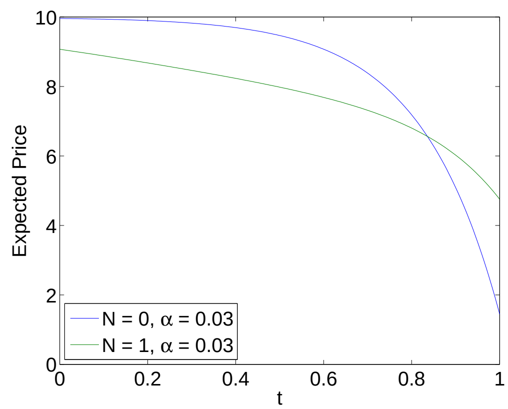

Figure 4.11: Depiction of the expected price (see Question IV.4).

α. It appears to be a concave function of α with a maximum roughly near $\alpha = 0.04$. Comparing this graph to Figures 4.7 and 4.8, we see that although the predator's expected revenue increases and the index tracker's expected revenue decreases as α gets smaller, the index tracker loses more than the predator gains for α small enough.

4.5.2 Effects Due to the Predator's Appearance

Question IV.4. *What is the effect of the predator's appearance?*

Figure 4.10 shows the index tracker's trading rate as a function of time for various values of α and N. For the values of α we considered, the entry of the predator causes the index tracker to sell more rapidly at small and large times and more slowly at intermediate times. We qualify this statement because for these price impact

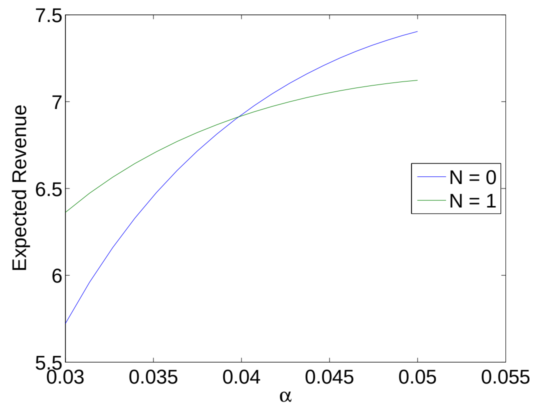

Figure 4.12: Depiction of the index tracker's expected revenue (see Question IV.4).

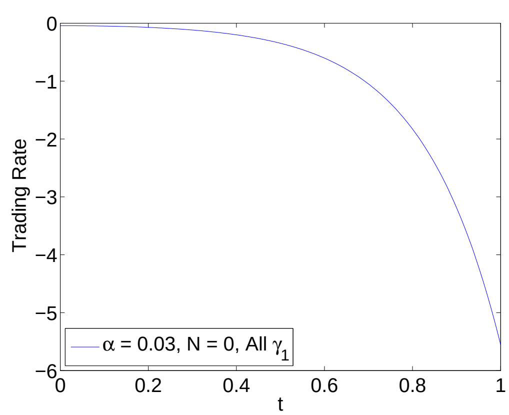

Figure 4.13: Depiction of the index tracker's trading rate (see Question IV.5).

parameters, the index tracker (distressed trader) sells more slowly for small times and more rapidly for large times in the one-period model of [69] and [216]. We did not attempt to find the smallest value of α for which this change in behavoir occurs. Also, while the index tracker's trading rate is decreasing when the predator is absent, it is only eventually decreasing after the predator's appearance.

In Figure 4.11, we have the expected price process as a function of t both in the predator's absence and presence. For the α we considered, the expected price process when the predator is in the market is initially smaller but is eventually much higher than it is in the absence of the predator. The same is true in the one-period model of [69] and [216].

Figure 4.12 depicts the index tracker's expected revenue as a function of α both in the predator's absence and presence. The entry of the predator increases the index tracker's expected revenue for small α and decreases the index tracker's expected revenue for large α. Observe that transition occurs roughly near $\alpha = 0.04$. It would be interesting to study whether or not this is the same value of α which maximizes the sum of the expected revenues in Figure 4.9 and, if so, why this is the case. Recall that in the one-period model of [69] and [216], the index tracker (distressed trader) always loses expected revenue as a result of the predator's entrance to the market.

4.5.3 Effects Due to the Price Impact Parameters

We will now try to understand the influence of the price impact parameters. We will vary each parameter one at a time, while keeping the remaining parameters as they were previously. We remark that all of the properties that we just observed still hold.

Question IV.5. *How does varying γ_1 affect our solution?*

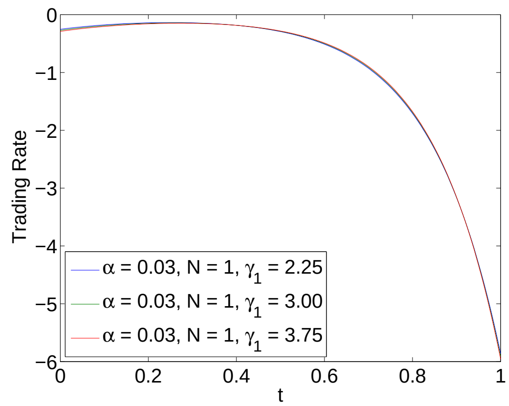

Figure 4.14: Depiction of the index tracker's trading rate (see Question IV.5).

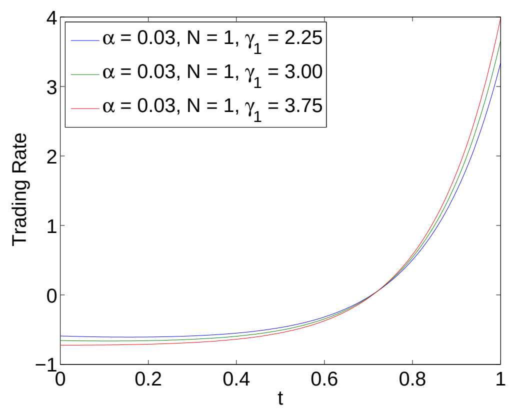

Figure 4.15: Depiction of the predator's trading rate (see Question IV.5).

Figure 4.13 shows the index tracker's trading rate as a function of t when the predator is not active. In this case, the index tracker's trading rate does not depend on γ_1. Recall that the distressed trader's trading rate in the predator's absence is also independent of the permanent price impact parameter in the one-period model of [69] and [216].

In Figure 4.14, we see the index tracker's trading rate as a function of t after the predator's appearance for various γ_1. The index tracker now responds to changes in γ_1. Specifically, the index tracker sells (slightly) more rapidly for small and large times and (slightly) more slowly for intermediate times as γ_1 increases.

Figure 4.15 depicts the predator's trading rate as a function of t for several values

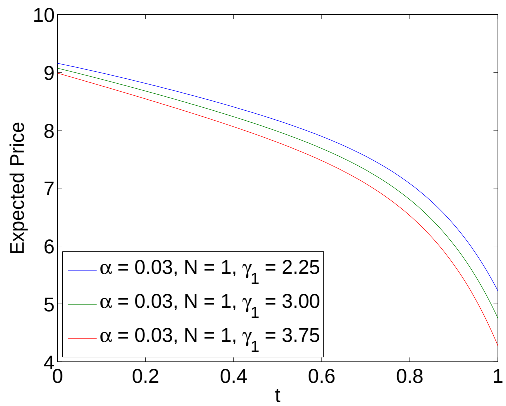

Figure 4.16: Depiction of the expected price (see Question IV.5).

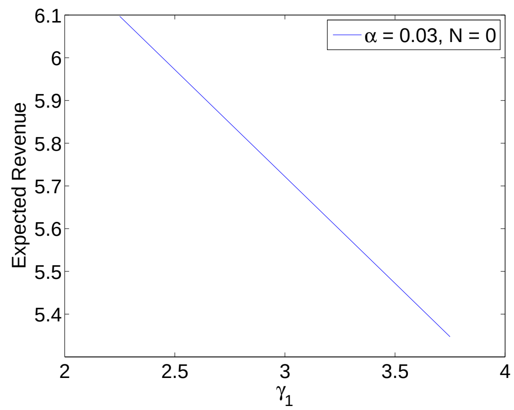

Figure 4.17: Depiction of the index tracker's expected revenue (see Question IV.5).

of γ_1. The predator sells more rapidly for small times and buys more rapidly for large times as γ_1 increases. The predator appears to sell for a slightly longer period of time as γ_1 increases. Comparing this graph to Figure 4.14, the predator's trading rate appears to be more sensitive to γ_1 than the index tracker's trading rate.

Figure 4.16 shows the expected price process as a function of t when the predator is active for various γ_1. It appears that the expected price process decreases for all times as γ_1 increases and that the magnitude of the drop increases with time. Apparently the upward price pressure due to the predator's increased trading rate for large times is outweighed by the downward price pressure due to the index tracker's liquidation (and the predator's short-selling) as γ_1 increases.

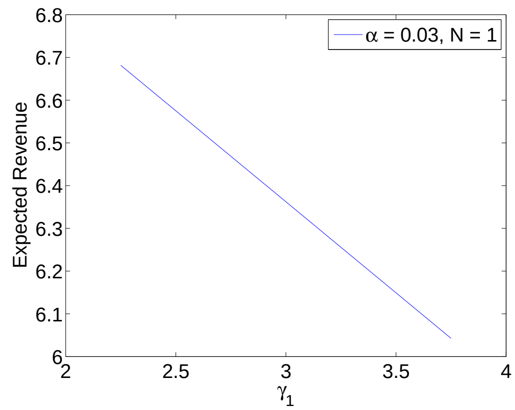

Figure 4.18: Depiction of the index tracker's expected revenue (see Question IV.5).

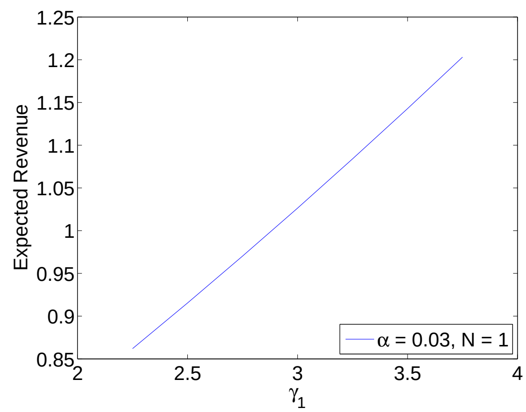

Figure 4.19: Depiction of the predator's expected revenue (see Question IV.5).

In Figure 4.17, we have the index tracker's expected revenue as a function of γ_1 when the predator is not active. This quantity decreases linearly with γ_1, which is also clear from Figure 4.13, since the index tracker's trading rate does not depend on γ_1.

Figure 4.18 depicts the index tracker's expected revenue as a function of γ_1 when the predator is active. After a close inspection, it appears that this is a decreasing concave function. The fact that this function is decreasing seems clear given Figures 4.14 and 4.16, since the index tracker's trading rate seems to only minimally change and the expected price process noticeably drops as γ_1 increases.

Figure 4.19 shows the predator's expected revenue as a function of γ_1. It appears

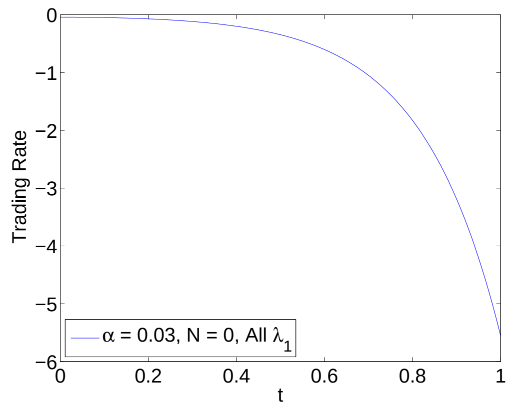

Figure 4.20: Depiction of the index tracker's trading rate (see Question IV.6).

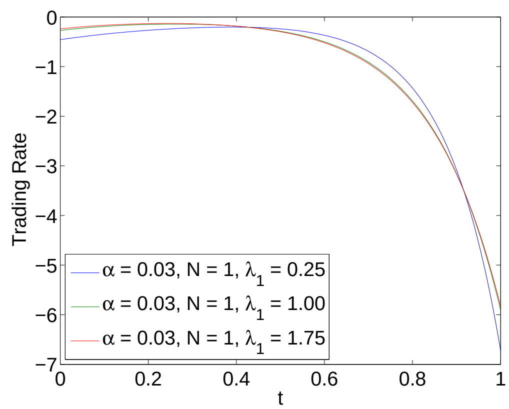

Figure 4.21: Depiction of the index tracker's trading rate (see Question IV.6).

to be an increasing convex function γ_1. Considering Figures 4.15 and 4.16, one might have guessed that this function is increasing: the predator is initially short-selling more rapidly and the drop in the expected price process grows with time as γ_1 increases.

Question IV.6. *How does varying λ_1 affect our solution?*

In Figure 4.20, we have the index tracker's trading rate as a function of t when the predator is absent from the market. This trading rate does not depend on λ_1, so the index tracker's trading rate does not depend on either of the price impact parameters in the predator's absence (see Figure 4.13). Again, the same is true of the distressed trader's trading rate in the one-period model of [69] and [216].

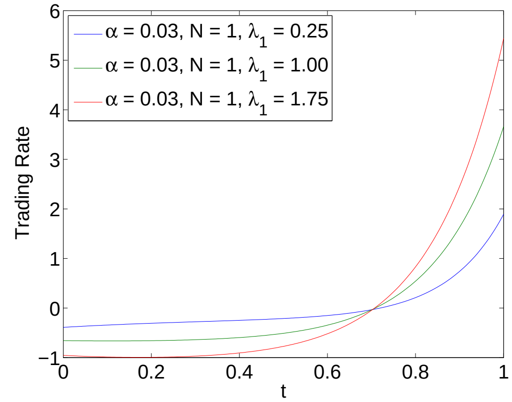

Figure 4.22: Depiction of the predator's trading rate (see Question IV.6).

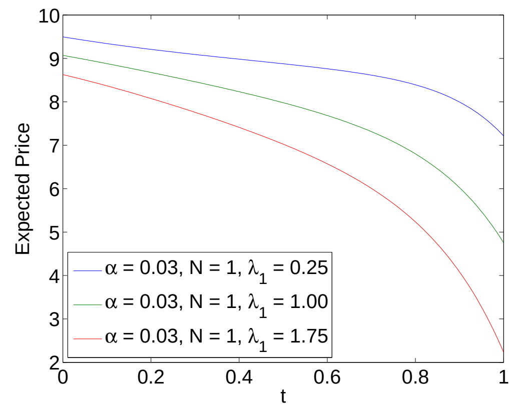

Figure 4.23: Depiction of the expected price (see Question IV.6).

Figure 4.21 depicts the index tracker's trading rate as a function of t after the predator's apperance. It appears that the index tracker sells more slowly for small and large times and more rapidly for intermediate times as λ_1 increases. Recall that this is the opposite of what we observe as γ_1 increases in Figure 4.14. Comparing these two figures, it also appears that the index tracker's trading rate might be more sensitive to λ_1 than γ_1.

Figure 4.22 shows the predator's trading rate as a function of t for several values of λ_1. The predator sells more rapidly for small times and buys more rapidly for large times as λ_1 increases. This is the same effect that we observe in Figure 4.15 as γ_1 increases, although it appears that the predator's trading rate may be more

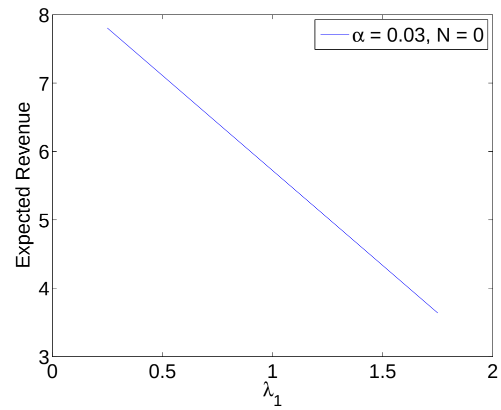

Figure 4.24: Depiction of the index tracker's expected revenue (see Question IV.6).

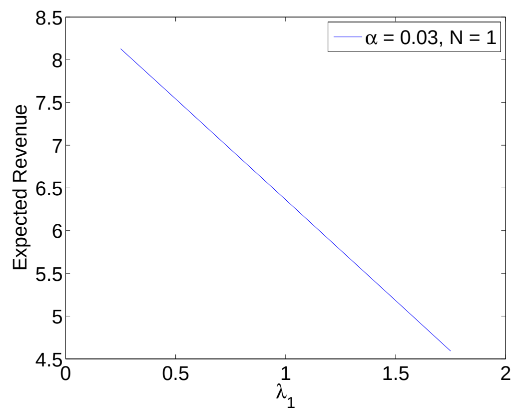

Figure 4.25: Depiction of the index tracker's expected revenue (see Question IV.6).

sensitive to λ_1 than to γ_1.

In Figure 4.23, we have the expected price process as a function of time when the predator is active for various λ_1. It appears that this function decreases for all times as λ_1 increases and that the magnitude of the drop increases with time. We observe the same phenomenon as γ_1 increases in Figure 4.16. Also, it appears that the concavity of the expected price process may change for small times as λ_1 decreases.

Figure 4.24 depicts the index tracker's expected revenue as function of λ_1 when the predator is not in the market. The expected revenue decreases linearly with λ_1. As we observed with γ_1 in Figure 4.17, the linear decrease is clear from Figure 4.20,

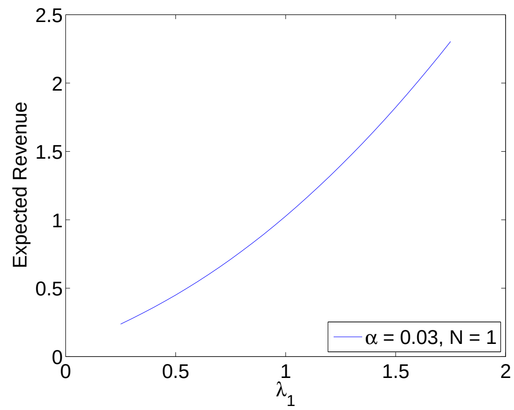

Figure 4.26: Depiction of the predator's expected revenue (see Question IV.6).

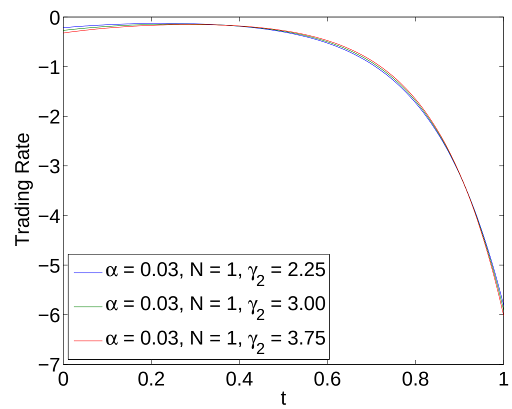

Figure 4.27: Depiction of the index tracker's trading rate (see Question IV.7).

since the index tracker's trading rate does not depend on λ_1.

Figure 4.25 shows the index tracker's expected revenue as function of λ_1 when the predator is present. In this case, a close inspection suggests that the index tracker's expected revenue is a decreasing concave function of λ_1. Recall that we observe the same thing when γ_1 increases in Figure 4.18. Also, notice that as λ_1 increases from 0.25 to 1.75, the index tracker loses much more in expected revenue when there is no predator than when the predator is active.

In Figure 4.26, we have the predator's expected revenue as a function of λ_1. It appears to be an increasing convex function. While we observed that the predator's expected revenue is also an increasing convex function of γ_1 in Figure 4.19, a comparison of these two figures suggests that it might be more sensitive to λ_1.

Question IV.7. *How does varying γ_2 affect our solution?*

Figure 4.27 depicts the index tracker's trading rate as a function of t for several values of γ_2. The index tracker sells (slightly) more rapidly for small and large times and (slightly) more slowly for intermediate times as γ_2 increases. Recall that we observe the same effect as γ_1 increases in Figure 4.14 and the opposite effect as λ_1 increases in Figure 4.21. Also, changing γ_2 appears to have the greatest effect on the index tracker's trading rate for very small times and very large times.

Figure 4.28 shows the predator's trading rate as a function of t for various γ_2. The predator sells (slightly) more slowly for small times and buys (slightly) more rapidly for large times as γ_2 increases. As λ_1 and γ_1 increase, we see the same phenomenon for large times; however, increasing these parameters has the opposite effect on the predator's trading rate for small times (see Figures 4.15 and 4.22). Note that the predator's trading rate seems to be less sensitive to γ_2 than it is to γ_1 and λ_1.

In Figure 4.29, we have the expected price process as a function of t for a few values of γ_2. This process decreases for all times as γ_2 increases, and the magnitude of the drop is the largest for intermediate times and fairly negligible for small and large times. While we also observe that the expected price process decreases as γ_1 and λ_1 increase, the magnitude of the drop appears to increase with time in those cases (see Figures 4.16 and 4.23).

Figure 4.30 depicts the index tracker's expected revenue as a function of γ_2. A close inspection suggests that it is a decreasing convex function of γ_2. From Figures 4.18 and 4.25, we observe that while the expected revenue still decreases with γ_1 and λ_1, it does so in a concave manner in those cases. The index tracker's expected revenue also appears to be less sensitive to γ_2 than to either γ_1 or λ_1.

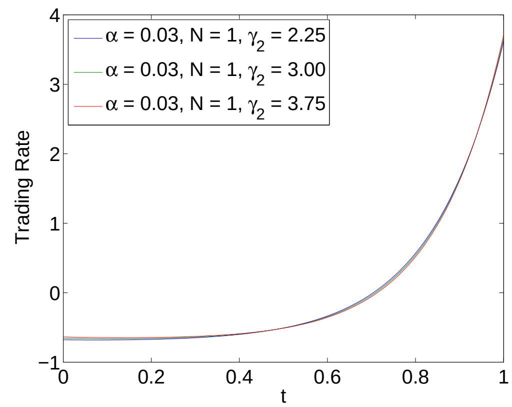

Figure 4.28: Depiction of the predator's trading rate (see Question IV.7).

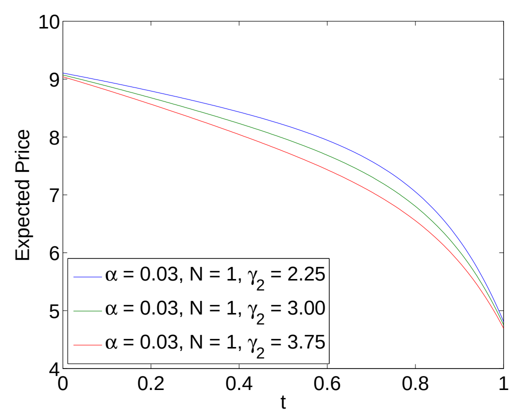

Figure 4.29: Depiction of the expected price (see Question IV.7).

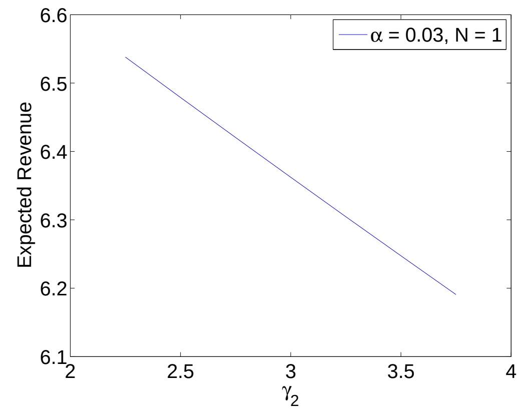

Figure 4.30: Depiction of the index tracker's expected revenue (see Question IV.7).

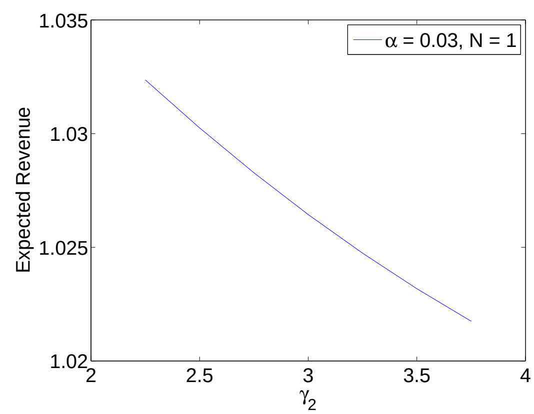

Figure 4.31: Depiction of the predator's expected revenue (see Question IV.7).

Figure 4.31 shows the predator's expected revenue as a function of γ_2. It appears to be a decreasing convex function of γ_2. While the predator's expected revenue is also a convex function of γ_1 and λ_1, it is an increasing function in each of those cases (see Figures 4.19 and 4.26). Still, the fact that the expected revenue is decreasing here is not too surprising given Figures 4.28 and 4.29, since the predator's trading rate changes minimally but the expected price process drops as γ_2 increases.

Question IV.8. *How does varying λ_2 affect our solution?*

In Figure 4.32, we have the index tracker's trading rate as a function of t for various λ_2. The index tracker sells more slowly for small and large times but more rapidly for intermediate times as λ_2 increases. Recall that we observe exactly the

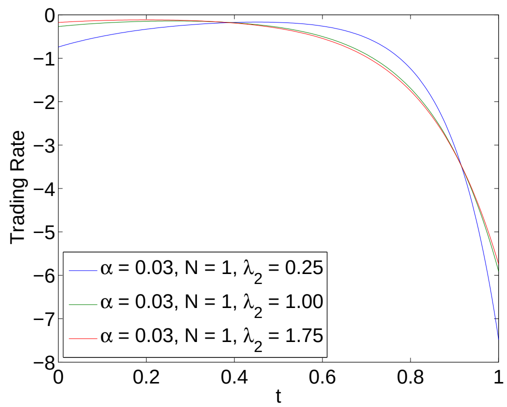

Figure 4.32: Depiction of the index tracker's trading rate (see Question IV.8).

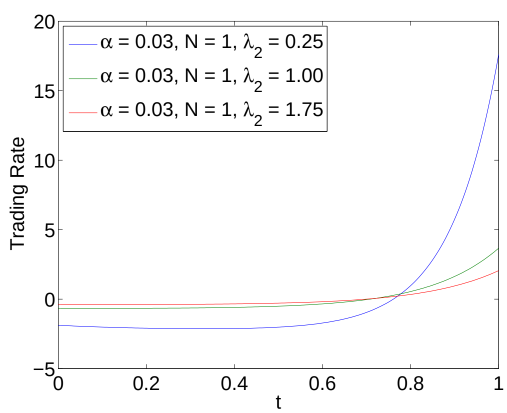

Figure 4.33: Depiction of the predator's trading rate (see Question IV.8).

same effects when λ_1 increases but the opposite effects when γ_1 or γ_2 increase (see Figures 4.14, 4.21, and 4.27). In general, it appears that the index tracker's trading rate is more sensitive to changes in the temporary price impact parameters than to changes in the permanent price impact parameters.

Figure 4.33 depicts the predator's trading rate as a function of t for several values of λ_2. The predator sells more slowly for small times and buys more slowly for large times as λ_2 increases. This is exactly the opposite relationship that the predator's trading rate has with γ_1 and λ_1 (see Figures 4.15 and 4.22). We see the same effect for small times but the opposite effect for large times when γ_2 increases. As is the case with the index tracker, it seems that the predator's trading rate is more sensitive overall to changes in the temporary price impact parameters than changes in the permanent price impact parameters.

Figure 4.34 shows the expected price process as a function of t for a few values of λ_2. The expected price process increases for all times as λ_2 increases. For large times, the concavity changes as λ_2 increases. Compared to the changes we observe when increasing γ_1, λ_1, and γ_2 in Figures 4.16, 4.23, and 4.29, these observations are somewhat surprising. For instance, in all other cases, the expected price process drops as the price impact parameter increases. Also, we did not notice any concavity change in the expected price process when varying γ_1 or γ_2. The concavity change we notice while varying λ_1 occurs for small times and is much less pronounced.

In Figure 4.35, we have the index tracker's expected revenue as a function of λ_2. It appears to be an increasing concave function. This is the only instance when we observe that the index tracker's expected revenue increases with a price impact parameter (see Figures 4.18, 4.25, and 4.30), although this observation is not too surprising given our observations about the expected price process in Figure

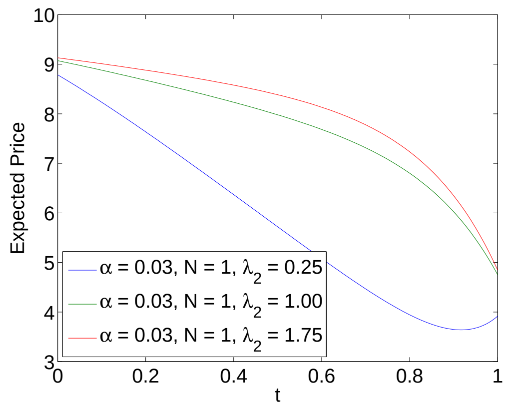

Figure 4.34: Depiction of the expected price (see Question IV.8).

4.34. Generally, it seems that the index tracker's expected revenue is much more sensitive to the temporary price impact parameters than to the permanent price impact parameters.

Figure 4.36 depicts the predator's expected revenue as a function of λ_2. It appears to be a decreasing convex function of λ_2. When varying any of the other price impact parameters, the predator's expected revenue has always appeared to be convex as well (see Figures 4.19, 4.26, and 4.31). The expected revenue seems to decrease when the predator's price impact parameters increase and increase when the index tracker's price impact parameters increase. As in the case of the index tracker, the predator's expected revenue appears to be much more sensitive to the temporary price impact parameters than to the permanent price impact parameters.

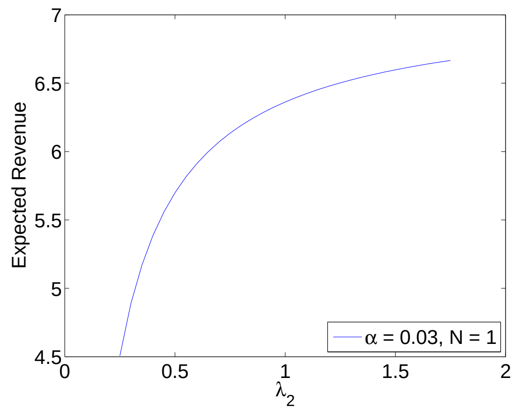

Figure 4.35: Depiction of the index tracker's expected revenue (see Question IV.8).

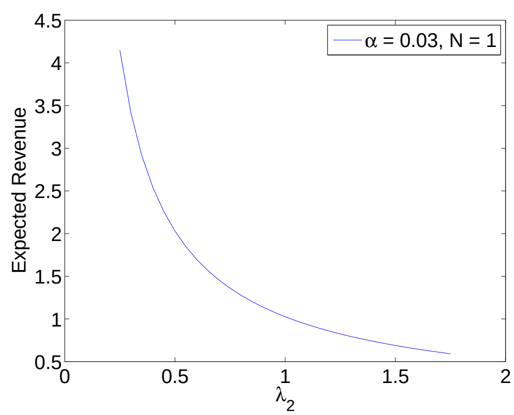

Figure 4.36: Depiction of the predator's expected revenue (see Question IV.8).

CHAPTER V

High-Roller Impact: A Large Generalized Game Model of Parimutuel Wagering

5.1 Introduction

Suppose that a collection of bettors are wagering on an upcoming event. The payoffs are determined via a *(frictionless) parimutuel system* if whenever Outcome i occurs, Bettor A receives

$$(\text{Total Amount Wagered})\left(\frac{\text{Bettor } A\text{'s Wager on Outcome } i}{\text{Total Amount Wagered on Outcome } i}\right).$$

The idea is that players with correct predictions will proportionally share the final betting pool. Prizes are reduced in practice by transaction costs such as the *house take*, a percentage fee collected by the *betting organizer* (or *house*). For example, Bettor A might only win

$$\kappa\,(\text{Total Amount Wagered})\left(\frac{\text{Bettor } A\text{'s Wager on Outcome } i}{\text{Total Amount Wagered on Outcome } i}\right) \tag{5.1}$$

when Outcome i occurs, if the house take is $(1-\kappa)\,\%$ for some $0<\kappa<1$.

This mechanism was invented in the context of horse race gambling by Oller in the late 1800's ([68]) and remains widely employed in that setting: In 2014, worldwide parimutuel betting on horse races totaled around seventy-five billion euros ([7]). It also typically determines wagering payoffs for other sports such as jai alai and races involving bicycles, motorcycles, motorboats, and greyhounds ([38]). Certain prizes for major lotteries such as Mega Millions, Powerball, and "EuroMillions" are computed in a parimutuel fashion ([6]). Parimutuel systems are increasingly popular methods for distributing payoffs in online prediction markets as well ([197]). Goldman Sachs Group, Inc., Deutsche Bank AG, CME Group Inc., Deutsche Börse AG, and ICAP have even facilitated the development of parimutuel derivatives on economic indicators ([38]).

The first scholarly publication on parimutuel wagering was written by Borel in 1938 ([57]), and more recent surveys and anthologies ([228]; [122]; [121]) attest to the substantial academic interest garnered by this topic since. A vast range of issues from optimal betting ([138]; [209]; [56]) to market efficiency ([123]; [25]; [134]) to market microstructure ([154]; [197]; [196]) has been extensively studied. Significant attention has been paid to strategic interactions among bettors ([205]; [184]; [151]; [227]; [201]).

Similar to the rise of high-frequency and algorithmic traders in financial markets, a growing number of parimutuel wagering event participants are organizations employing large-scale strategies based upon advanced mathematical, statistical, and computational techniques ([145]). There are fundamental differences between these bettors and more traditional wagerers. The new firms typically have access to vast budgets, making their betting totals orders of magnitude beyond the amounts wagered by regular players. Often, they can place their wagers at speeds impossible for ordinary bettors to match. Presumably, their use of complex methods also makes their forecasts and corresponding wagering strategies generally superior.

The house collects a percentage of the total amount wagered and, therefore, may initially benefit from the presence of large-scale wagering firms. After all, their activities should increase the size of the pool, at first anyway. The factors just described are thought to put ordinary bettors at an extreme disadvantage, though. Since payouts are calculated according to (5.1), ordinary bettors' profits may even directly decline as a result of the large-scale firms' wagers. If this discourages enough regular players from betting, then pool sizes may eventually dwindle, hurting the house's revenue. In fact, many betting organizers have publicly expressed strong concerns about the new breed of wagerers. Betting organizers have even occasionally banned these participants from parimutuel wagering events ([145]).

How reasonable is this narrative? Our goal is to quantify the impact of large-scale participants in parimutuel wagering events on the house and ordinary bettors.

First, using the theory of *large generalized games*, i.e., games with a continuum of diffuse (or non-atomic/minor) players and finitely many atomic (or major) players, we develop a model of parimutuel betting. The bets made by individual atomic players affect all others because they change the final payoff per unit bet on Outcome i:

$$\kappa\left(\frac{\text{Total Amount Wagered}}{\text{Total Amount Wagered on Outcome } i}\right). \tag{5.2}$$

Aggregate decisions made by the diffuse players also affect every player for the same reason. A key feature is that an individual diffuse player cannot change (5.2) by revising her wagers. In fact, her specific choices have no effect whatsoever on the rest of the game's participants.

We view diffuse and atomic players as stand-ins for ordinary bettors and large-scale wagering firms, respectively. The approximation is motivated by the observation that the total amount wagered by a single traditional bettor is generally negligible compared to the total amount wagered by an entire betting firm.

Our main theoretical result, Theorem V.7, provides necessary and sufficient conditions for the existence and uniqueness of a pure-strategy Nash equilibrium. Other scholars have shown the existence of equilibria in a broad class of large generalized games ([32]; [33]; [70]; [206]). Such results often rely upon sophisticated technology including variants of the Kakutani fixed-point theorem. We choose to employ a more elementary fixed-point argument instead. Advantages of our approach include its simplicity and the possibility of proving the equilibrium's uniqueness. A simple algorithm for computing relevant equilibrium quantities immediately presents itself as well.

Having such an algorithm allows us to analyze our problem in specific scenarios. In accordance with the prevailing narrative, we observe that because of the atomic player, the house is temporarily better off and the diffuse players are worse off in Example V.12. For varying reasons, at least one of these effects is not observed in each of the remaining situations.

In Example V.13, the diffuse players are better off in the presence of the atomic player. Intuitively, when the event is *too close to call*, the atomic player can bet on the *wrong* outcome even if her prediction is assumed to be quite accurate. Such an error is to the advantage of the diffuse players.

In Example V.14, the diffuse players *believe* that they are better off when there is an atomic player. Roughly, if the diffuse players' beliefs are too homogeneous but the atomic player disagrees with them, the diffuse players' expected profits per unit bet rise when the atomic player takes the other side of their wagers.

In Example V.15, we argue that the diffuse players are better off but the house is (immediately) worse off because of the atomic player, exactly the opposite of the prevailing narrative.[1] To make this point, we recast our model as a two-stage game taking into account the house's strategic decisions. Effectively, because the atomic

[1]Recall that the alleged decline in the house's revenue is thought to occur over time.

player considers her impact on (5.2), she has a lower tolerance for unfavorable betting conditions than the diffuse player. When she is absent, this means that the house can more easily prey upon the diffuse players.

Before offering further details, we more thoroughly discuss related literature in Section 5.2. We carefully present our model in Section 5.3 and our main theoretical result in Section 5.4. We numerically investigate our concrete examples in Section 5.5. Section 5.6 highlights a few technical aspects of Section 5.4's work. We give our longer formal proofs in Appendices 5.7 and 5.8.

5.2 Related Literature

The individual states of our diffuse players are only coupled via the empirical distribution of controls. Since each diffuse player is too small to influence this distribution, she treats it as fixed when determining her own strategy. Assumptions like these have appeared in the literature on continuum games ([27]; [214]; [168]; [207]) and mean-field games ([158]; [133]).[2]

Our paper bears a stronger resemblance to work in the former category. For instance, we model parimutuel wagering as a static game. Such a choice is quite common in a continuum game study; however, stochastic differential games are more often the focus in mean-field game theory. Also, we restrict ourselves to an intuitive argument for viewing ordinary bettors as diffuse players. Papers on mean-field games often rigorously present their continuum model as a limit of finite population models, while those on continuum games typically emphasize other issues.

General mean-field interactions among players can be described by complex functions of the empirical distributions of states and/or controls. On the other hand,

[2] Originally, players were coupled via the empirical distribution of states, not controls, in the mean-field game literature. Recent advances have shown that models with additional interactions through the controls also have promising applications, say in the contexts of price impact, optimal execution, high-frequency trading, and oligopolistic energy market problems ([117]; [71]; [115]; [81]; [80]).

parimutuel wagerers affect one another through (5.2) alone, a comparatively simple scenario. This makes parimutuel wagering an especially strong candidate for modeling by either theory. Continuum games have already been applied in this way ([182]; [234]).

Watanabe considered a two-stage game with a betting organizer and a continuum of risk-neutral diffuse players with heterogeneous beliefs ([234]). First, the betting organizer selects a value for the house take. The diffuse players then determine whether to place a unit bet on one of two outcomes or bet nothing at all. Using techniques from set-valued analysis, Watanabe showed that an equilibrium always exists, provided the house take is not too large. As long as each player can only bet negligible amounts, Watanabe also found that equilibria in parimutuel wagering games are *regular*. That is, if a player wagering on Outcome i believes that Outcome i will occur with probability p, all players who believe that Outcome i will occur with probability $p' > p$ also wager on Outcome i. The paper's results on the betting organizer's optimal strategy were in the context of specific examples.

Ottaviani and Sørenson developed their continuum game to explain two phenomena frequently observed in the context of parimutuel wagering on horse races: *late informed betting* and the *favorite-longshot bias* ([182]). The first states that more accurate information about a race's outcome can be gleaned from late bets than early bets. The second says that the public tends to excessively bet on unlikely outcomes and wager too little on likely outcomes. A continuum of privately informed risk-neutral players decide when to place their individual bets in a discrete-time setting. They can wager a unit amount on one of two outcomes or abstain from betting. The paper gave conditions under which all of these players simultaneously wager at the terminal time, and the corresponding equilibrium is shown to always feature the favorite-longshot bias.

Others have more implicitly created infinite-player parimutuel wagering models by assuming that there are *many* bettors ([134]; [183]). These references have sought to identify other possible causes of the favorite-longshot bias.

The most important new feature of our setup is that, in addition to the diffuse bettors, we introduce an atomic bettor. The parimutuel wagering studies we just discussed only incorporate diffuse players. We would not be able to understand the effects of large-scale wagering organizations on ordinary bettors, if we made a similar assumption.

Because of this addition, our model belongs to the class of continuum game models known as *large generalized games*. Games that include both diffuse and atomic players can be found in mean-field game theory as well under the heading *major-minor player models* ([132]; [174]; [175]; [141]; [232]). Applications of large generalized games are known to be diverse and already include a collection of problems from politics ([90]) to oligopolistic markets ([238]). General results on the existence of equilibria in large generalized games have also been obtained ([32]; [33]; [70]; [206]). We choose not to rely upon these, as the simple structural aspects of parimutuel wagering just discussed, combined with a convenient modeling assumption (see Section 5.3), allow us to use elementary arguments.

Ultimately, we employ a mean-field approximation for the standard reason: By doing so, we make our model tractable, hopefully while preserving the critical macroscopic properties of our original problem. Issues other than the impact of large-scale wagering organizations have been resolved in finite-player settings, though such studies have usually invoked other strong assumptions.

Weber gave sufficient conditions for the existence of an equilibrium in a simultaneous parimutuel betting game with N atomic players, each of whom is risk-neutral and must bet a specific total amount ([237]). Watanabe, Nonoyama, and Mori's setup and goals are similar to those in Watanabe's continuum game paper, except that the diffuse players in the latter are replaced by finitely many atomic players ([235]; [234]). Explanations of the favorite-longshot bias have been offered using equilibrium results for N-player parimutuel wagering games ([77]; [151]; [184]; [204]). Games in which finitely many bettors wager sequentially have also been investigated for various purposes ([103]; [85]; [152]; [151]; [229]). For example, Thrall showed that if risk-neutral atomic bettors with homogeneous beliefs wager sequentially, their profits tend to zero as the number of bettors increases ([229]). Note that some of these works do consider limiting cases in which the population of wagerers grows arbitrarily large to complement their other insights ([229]; [184]).

5.3 Model Details

Our players have the opportunity to wager on an event that can unfold in two mutually exclusive ways: Outcome 1 might occur. If not, Outcome 2 will occur. For now, we view $\kappa \in (0, 1)$ to be exogenously given. Inspired by Watanabe et al. ([235]; [234]), we later informally consider what happens when we allow the house take to be optimally selected by the betting organizer in the first stage of a two-stage game (see Example V.15). Our results in Section 5.4 are unaffected by such a shift.

The unit interval describes the diffuse bettors' views on the likelihood of Outcome 1: the bettors whose views are indexed by $p \in [0, 1]$ believes that Outcome 1 will occur with probability p. Initially, each diffuse bettor has some (negligible) unit wealth. A finite Borel measure μ with a continuous everywhere positive density characterizes the distribution of the diffuse bettors. More precisely, the total initial wealth of all diffuse bettors whose views are contained in a Borel set A is $\mu(A)$.

That μ has a continuous everywhere positive density is our *convenient modeling assumption* from Section 5.2. It is similar to a key hypothesis in Ottaviani and Sørenson's work, although the posterior beliefs of their diffuse bettors are obtained after updating a common prior belief using a private signal and Bayes' rule ([182]). Effectively, Watanabe assumed that μ need not have a density, and even when it does, the density need not be positive everywhere ([234]). These choices necessitated a set-valued approach, which we are able to avoid.

Continuity merely simplifies a few of our arguments, e.g., see Step 6 of Theorem V.7's proof. The rest of the assumption plays a more critical role. Results on the existence of equilibria in parimutuel wagering games often include a hypothesis such as the following: for any given outcome, at least two bettors believe that the outcome will occur with positive probability.[3] By supposing that the density is positive, we assume this as well. Watanabe has shown that an equilibrium may not be unique, if $\mu(\{p\}) > 0$ for some fixed p ([234]). Obviously, this situation is prevented by the density's existence.

Our atomic player believes that Outcome 1 will occur with probability $q \in [0,1]$. She has (non-negligible) finite initial wealth $w > 0$.

Throughout, we treat all players' beliefs as exogenously determined. We do not address how the players generate their estimates; however, this process is of great interest both practically and academically ([122]; [121]). For our theoretical results in Section 5.4, we also do not specify how the players' estimates compare to the *actual* probability that Outcome 1 will occur. Since large-scale betting organization allegedly produce highly accurate forecasts, we could choose q to be some small

[3] This is true of Weber's work, for instance ([237]). Roughly, if only one player believes that a particular outcome will occur with positive probability, she should wager an arbitrarily small amount on that outcome. In the absence of a positive minimum bet size constraint, it follows that an equilibrium does not exist.

perturbation of the actual probability of Outcome 1. We informally experiment with this extra assumption in Examples V.12 and V.13.

All players decide how much to wager on each outcome. Their choices are constrained only by their initial wealth: a betting strategy is *feasible* (or *admissible*) for an individual bettor as long as the sum of her wagers is no more than her wealth. For example, a bettor could choose to wager 100% of her wealth on Outcome 1, 55% of her wealth on Outcome 1 and 30% of her wealth on Outcome 2, or not wager at all. We formalize this as follows.[4]

Definition V.1. A *feasible strategy profile for the diffuse players* is a measurable function

$$f = (f_1, f_2) : (0,1) \longrightarrow \left\{(x_1, x_2) \in \mathbb{R}^2_{\geq 0} \, : \, x_1 + x_2 \leq 1\right\}.$$

A *feasible strategy profile for the atomic player* is a vector $a = (a_1, a_2) \in \mathbb{R}^2_{\geq 0}$ such that

$$a_1 + a_2 \leq w.$$

We call the pair (f, a) a *feasible strategy profile.*

Under (f, a), the atomic player wagers a_i on Outcome i. Each diffuse player who believes that Outcome 1 will occur with probability p wagers $f_i(p) \times 100\%$ of her (negligible) unit initial wealth on Outcome i.

Our space of feasible strategy profiles is slightly atypical. Previously, diffuse players in parimutuel wagering games have only been able to place unit bets, if they bet at all ([182]; [234]). We could have made this restriction as well without loss of generality due to Proposition V.5. Watanabe allowed groups of diffuse players to wager differently, even if they held identical beliefs ([234]). In equilibrium, such

[4] In practice, concerns about large-scale wagering firms are also driven by their alleged ability to bet *faster* than ordinary players. Our static game only has two outcomes, so it does not make sense to model this feature here. We hope to revisit the issue in a future work.

a discrepancy could only arise among the diffuse players who believed that their expected profits would be zero. We encounter a related ambiguity in our framework (see our discussion of Proposition V.5). For us, the μ-measure of this set of bettors is zero, and we anticipate that all of our main results would remain the same, if we were to relax the diffuse bettors' *same beliefs-same bets* restriction. More significantly, atomic players have been frequently constrained to wager a fixed amount in total or a unit amount on a single outcome when they bet ([237]; [235]; [77]; [184]).[5] Proposition V.6 suggests that imposing these restrictions would have a severe effect.

The total amount that the diffuse players wager on Outcome i, denoted d_i, is given by

$$d_i = \int_0^1 f_i\left(p\right)\mu\left(dp\right).$$

Since μ has a density, we immediately confirm that the bets placed by any given diffuse player are too small to affect the amount wagered on any specific outcome. Of course, if she revises her strategy, then a particular diffuse player affects neither the total amount wagered nor (5.2). All of these quantities could change when aggregations of diffuse players, that is, collections of diffuse players whose beliefs are contained in a Borel set A with positive μ-measure, revise their wagers.

Payoffs are determined according to (5.1). Each player selects her wagering strategy simultaneously in order to maximize her expected profit according to her beliefs. We implicitly assume that every bettor knows κ, μ, q, and w, so that she can select the best response to her opponents' collective actions. Similar assumptions can be found in many other equilibrium studies on parimutuel wagering ([237]; [235]; [234]; [77]; [184]).[6]

[5] An exception is Cheung's thesis, though that work's focus is quite different from our own ([85]).

[6] Even so, our players might seem unrealistically knowledgeable and confident in their beliefs. Real bettors would presumably have a more complex prior for the outcomes' likelihoods. They probably would not have access to such comprehensive information about their opponents. If they somehow had a sense of these details, they might also wish to update their own forecasts.

Since each diffuse player starts out with negligible unit wealth, technically, we should only discuss the expected profits of diffuse bettors whose views lie in some Borel set A. We nevertheless compute and refer to the expected profits of an individual diffuse bettor in an obvious, but admittedly informal, way. Doing so helps motivate our definition of a *pure-strategy Nash equilibrium* (see Definition V.3) and highlight the intuition underlying our results.

A seemingly more formidable concern is how to handle (5.1) in pathological cases. It is trivial to produce a feasible strategy profile (f, a) such that for some p, we have $f_j(p) > 0$ but

$$d_j = a_j = 0.$$

Naïvely translating (5.1), we conclude that a diffuse player whose views are indexed by p receives

$$\kappa \left(\sum_{i=1}^{2} (d_i + a_i) \right) \left(\frac{f_j(p)}{d_j + a_j} \right) \tag{5.3}$$

whenever Outcome j occurs.

If

$$d_1 = d_2 = a_1 = a_2 = 0,$$

then the total amount wagered is zero. Practically, the betting organizer would probably cancel such an event, which suggests that it is natural to set all players' payoffs to zero in this scenario.

An ad-hoc but simple way to address these objections could be to consider κ as some (publicly known) negative perturbation of the true κ (abusing notation). The idea is that each player would model transaction costs as being higher than their actual value, artificially lowering their perceived edge and encouraging them to bet more cautiously than they would otherwise.

A more thorough study could begin by determining whether or not approximating parimutuel wagering using the Nash equilibrium solution concept is, indeed, reasonable. If it is, one might endow the bettors with more sophisticated priors and enable them to update their beliefs using the equilibrium implied probabilities (see Definition V.2). This would lead to an extra condition in Definition V.3.

If the solution concept is unreasonable, one could devise a new scenario in which players independently determine their betting strategies according to individual reference models and appropriately penalized alternative models for outcome likelihoods, as well as their opponents' parameters. Broadly speaking, this treatment of a single player's optimization problem has seen widespread use across macroeconomics and finance ([124]). One might then investigate what unfolds when all players simultaneously participate in the same parimutuel wagering event.

We leave further consideration of these topics for a future work.

Alternatively, we might have

$$d_i + a_i > 0$$

for $i \neq j$. There are now two appealing options for the diffuse player's payoff. First, we might choose to set the payoff to zero. Practically, no bettor would receive a payout, if Outcome j occurred but no one wagered on it. We also might set the payoff to $+\infty$, as in Watanabe's work ([234]). In practice, if the amount wagered on Outcome j were zero, each player who believed that Outcome j would occur with positive probability would want to place an arbitrarily small bet on Outcome j. Setting the payoff to $+\infty$ captures this intuition.

We choose the first option, but selecting the second instead would not change our results. Only equilibrium payoffs need to be computed, and in an equilibrium, positive amounts are always wagered on both outcomes. Essentially, the scenario we describe never arises. One reason is that we do not allow the *trivial* (or *null*) equilibrium in which no player wagers. We could,[7] but as Watanabe observed, that case is comparatively uninteresting and practically unimportant ([234]). Roughly, the other reason is the same as our justification for possibly setting the payoff to $+\infty$.

Before making this discussion precise, we introduce some notation.

Definition V.2. Given a feasible strategy profile (f, a) such that at least one of the d_i's or a_i's is positive, the *implied (or subjective) probability* that Outcome 1 will occur, denoted $P^{f,a}$, is defined by

$$P^{f,a} = \frac{d_1 + a_1}{\sum_{i=1}^{2} (d_i + a_i)}.$$

[7] This situation could be considered an equilibrium because if any single player unilaterally revised her wagers, intuitively, she should incur a loss of at least $(1 - \kappa)\,\%$.

We refer to

$$1 - P^{f,a} = \frac{d_2 + a_2}{\sum_{i=1}^{2} (d_i + a_i)}$$

as the *implied (or subjective) probability* that Outcome 2 will occur.

$P^{f,a}$ is the ratio of the amount wagered on Outcome 1 to the total amount wagered, assuming the latter is positive. Our previous discussion implies that $P^{f,a} \in (0,1)$ in equilibrium.[8] Since the argument was informal, we do not yet take this as fact. In particular, the amount received by a diffuse player who believes that Outcome 1 will occur with probability p is[9]

$$\kappa \left(\sum_{i=1}^{2} (d_i + a_i) \right) \left(\frac{f_1(p) \mathbb{1}_{\{d_1 + a_1 \neq 0\}}}{d_1 + a_1} \right) = \frac{\kappa f_1(p) \mathbb{1}_{\{P^{f,a} \neq 0\}}}{P^{f,a}}$$

and

$$\kappa \left(\sum_{i=1}^{2} (d_i + a_i) \right) \left(\frac{f_2(p) \mathbb{1}_{\{d_2 + a_2 \neq 0\}}}{d_2 + a_2} \right) = \frac{\kappa f_2(p) \mathbb{1}_{\{P^{f,a} \neq 1\}}}{1 - P^{f,a}}$$

when Outcomes 1 and 2 occur, respectively. Hence, this diffuse player believes that her expected profit is

$$f_1(p) \left(\frac{\kappa \mathbb{1}_{\{P^{f,a} \neq 0\}} p}{P^{f,a}} - 1 \right) + f_2(p) \left(\frac{\kappa \mathbb{1}_{\{P^{f,a} \neq 1\}} (1-p)}{1 - P^{f,a}} - 1 \right). \tag{5.4}$$

Similarly, the atomic player thinks that her expected profit is

$$a_1 \left(\frac{\kappa \mathbb{1}_{\{P^{f,a} \neq 0\}} q}{P^{f,a}} - 1 \right) + a_2 \left(\frac{\kappa \mathbb{1}_{\{P^{f,a} \neq 1\}} (1-q)}{1 - P^{f,a}} - 1 \right). \tag{5.5}$$

Definition V.3. A *pure-strategy Nash equilibrium* is a feasible strategy profile $(f^\star, a^\star)$ such that

(i) at least one of the $d_i^\star$'s or $a_i^\star$'s is positive;

[8] We later observe that $P^{f,a} \in (1-\kappa, \kappa)$ in equilibrium (see Step 4 of Theorem V.7's proof).

[9] Of course, we still abuse notation here. When one of our indicator functions is equal to zero, the corresponding fraction is actually of the form $0/0$, not zero as we suppose.

(ii) for any $p \in [0,1]$,

$$f_1^\star(p)\left(\frac{\kappa\,\mathbb{1}_{\left\{P^{f^\star,a^\star}\neq 0\right\}}\,p}{P^{f^\star,a^\star}}-1\right)+f_2^\star(p)\left(\frac{\kappa\,\mathbb{1}_{\left\{P^{f^\star,a^\star}\neq 1\right\}}\,(1-p)}{1-P^{f^\star,a^\star}}-1\right)$$
$$=\sup_{\substack{b_1,b_2\geq 0\\ b_1+b_2\leq 1}}\left\{b_1\left(\frac{\kappa\,\mathbb{1}_{\left\{P^{f^\star,a^\star}\neq 0\right\}}\,p}{P^{f^\star,a^\star}}-1\right)+b_2\left(\frac{\kappa\,\mathbb{1}_{\left\{P^{f^\star,a^\star}\neq 1\right\}}\,(1-p)}{1-P^{f^\star,a^\star}}-1\right)\right\};$$

(iii) and

$$a_1^\star\left(\frac{\kappa\,\mathbb{1}_{\left\{P^{f^\star,a^\star}\neq 0\right\}}\,q}{P^{f^\star,a^\star}}-1\right)+a_2^\star\left(\frac{\kappa\,\mathbb{1}_{\left\{P^{f^\star,a^\star}\neq 1\right\}}\,(1-q)}{1-P^{f^\star,a^\star}}-1\right)$$
$$=\sup_{\substack{b_1,b_2\geq 0\\ b_1+b_2\leq w\\ d_1^\star,d_2^\star,b_1\text{ or }b_2>0}}\left\{b_1\left(\frac{\kappa\,\mathbb{1}_{\left\{P^{f^\star,b}\neq 0\right\}}\,q}{P^{f^\star,b}}-1\right)+b_2\left(\frac{\kappa\,\mathbb{1}_{\left\{P^{f^\star,b}\neq 1\right\}}\,(1-q)}{1-P^{f^\star,b}}-1\right)\right\}.$$

(i) formally excludes the case in which the total amount wagered is zero. (ii) and (iii) ensure that each player maximizes her expected profit according to her beliefs, given her opponents' wagers.[10]

First, observe that each player requires very little information about her opponents' strategies. For a given diffuse bettor, knowing $P^{f^\star,a^\star}$ alone is enough. The atomic player must be able to compute $P^{f^\star,b}$ for all of her feasible strategy profiles b, so it is sufficient for her to know $d_1^\star$ and $d_2^\star$. The difference for the two kinds of players reflects that an individual diffuse player cannot affect the implied probability of Outcome 1, while the atomic player can. These remarks explain why we informally claim that only the atomic player and aggregations of diffuse players affect the other participants. Notice that each player's strategy depends *anonymously* on her opponents' bets: how, specifically, her opponents wagers produced $P^{f^\star,a^\star}$, $d_1^\star$, and $d_2^\star$ is irrelevant.

Although we only optimize over b in (iii), the apparent possibility that $d_1^\star = d_2^\star = 0$

[10] In (ii), we determine the equilibrium wagers for all diffuse bettors whose beliefs are indexed by p; however, a single diffuse bettor with these beliefs solves the same maximization problem.

motivates our use of the extra constraint

$$d_1^\star, d_2^\star, b_1 \text{ or } b_2 > 0.$$

One might be concerned that we do not consider the feasible strategy $b_1 = b_2 = 0$ for the atomic player, if $d_1^\star = d_2^\star = 0$. Recall that all players receive a payoff of zero in such a situation. A simple calculation shows that the supremum is then also zero, so no issue is caused by our omission.

The last important concept for our modeling framework is uniqueness.

Definition V.4. A pure-strategy Nash equilibrium $(f^\star, a^\star)$ is *unique* if for any other pure-strategy Nash equilibrium $(f^\diamond, a^\diamond)$, we have $f^\star = f^\diamond$ μ-a.s. and $a^\star = a^\diamond$.

We allow $f^\star$ and $f^\diamond$ to disagree on a set of μ-measure zero since, ultimately, all relevant equilibrium quantities such as the implied probabilities are unaffected by such a difference. Soon, we see that $f^\star(p)$ and $f^\diamond(p)$ must be equal for all but two points, at most. There is only uncertainty about the behavior of the diffuse bettors who believe that their expected profits are zero (cf. our discussion about our space of feasible strategy profiles).

5.4 Theoretical Results

We now state and prove[11] our theoretical results, beginning with Propositions V.5 and V.6. The former describes how the diffuse players should wager in response to the atomic player's strategy. The latter tells us how the atomic player should bet, given the diffuse players' wagers. We use these observations to prove Theorem V.7, our main result.[12] Recall that it offers necessary and sufficient conditions for the existence and uniqueness of a pure-strategy Nash equilibrium.

[11] We describe the ideas underlying all of our proofs in Section 5.4; however, we delay our formal arguments for Proposition V.6 and Theorem V.7 until Appendices 5.7 and 5.8, respectively.

[12] Roughly, this suggests that our large generalized game can almost be viewed as a game with two players: the mean-field of diffuse players and the atomic player.

We conclude Section 5.4 with Corollaries V.8, V.9, and V.10. Corollary V.8 says that the atomic player wagers on a particular outcome if and only if the final expected profit per unit bet on that outcome is positive. The next corollary states that our equilibria are regular in a particular sense, while Corollary V.10 says that the implied probability of Outcome 1 tends to 0.5 uniformly as the house take approaches 50%.

Proposition V.5. *Let* (f, a) *be a feasible strategy profile such that* d_1*,* $d_2 > 0$*.* f *satisfies*

$$\begin{aligned} & f_1(p)\left(\frac{\kappa p}{P^{f,a}} - 1\right) + f_2(p)\left(\frac{\kappa\,(1-p)}{1-P^{f,a}} - 1\right) \\ & \quad = \sup_{\substack{b_1, b_2 \geq 0 \\ b_1 + b_2 \leq 1}} \left\{ b_1\left(\frac{\kappa p}{P^{f,a}} - 1\right) + b_2\left(\frac{\kappa\,(1-p)}{1-P^{f,a}} - 1\right)\right\} \end{aligned} \tag{5.6}$$

for all $p \in [0, 1]$ *if and only if*

$$f_1(p) = \begin{cases} 1 & \text{if } p > P^{f,a}/\kappa \\ & \\ 0 & \text{if } p < P^{f,a}/\kappa \end{cases} \qquad f_2(p) = \begin{cases} 1 & \text{if } 1-p > \left(1-P^{f,a}\right)/\kappa \\ & \\ 0 & \text{if } 1-p < \left(1-P^{f,a}\right)/\kappa \end{cases}.$$

First, notice that $P^{f,a} \in (0, 1)$ since both d_1 and d_2 are positive. Comparing (5.6) and (ii) of Definition V.3, we find that Proposition V.5 provides a simple characterization of the diffuse players' equilibrium strategies in this case, given the bets of the atomic player. Our assumption is not too restrictive, as Step 1 of Theorem V.7's proof says that d_1 and d_2 are always both positive in equilibrium.

Three distinct groups of diffuse bettors emerge: The first group, containing the diffuse bettors who believe that Outcome 1 will occur with probability greater than $P^{f,a}/\kappa$, wager all of their initial wealth on Outcome 1. Diffuse bettors who think that Outcome 2 will occur with probability greater than $\left(1 - P^{f,a}\right)/\kappa$ make up the second group. These players bet their entire fortunes on Outcome 2. The remaining diffuse players, except those whose beliefs are indexed by $p = P^{f,a}/\kappa$ or $\left(1 - P^{f,a}\right)/\kappa$, do not wager at all. None of the groups overlap, since $0 < \kappa < 1$ implies that

$$1 - \left(\frac{1 - P^{f,a}}{\kappa}\right) < \frac{P^{f,a}}{\kappa}. \tag{5.7}$$

The proof's underlying intuition is easy to explain. It is equivalent to show that f satisfies (5.6) on $[0,1]$ if and only if

$$f_1(p) = \begin{cases} 1 & \text{if } \frac{\kappa p}{P^{f,a}} - 1 > 0 \\ \\ 0 & \text{if } \frac{\kappa p}{P^{f,a}} - 1 < 0 \end{cases} \qquad f_2(p) = \begin{cases} 1 & \text{if } \frac{\kappa(1-p)}{1-P^{f,a}} - 1 > 0 \\ \\ 0 & \text{if } \frac{\kappa(1-p)}{1-P^{f,a}} - 1 < 0 \end{cases}.$$

The terms

$$\frac{\kappa p}{P^{f,a}} - 1 \qquad \text{and} \qquad \frac{\kappa(1-p)}{1 - P^{f,a}} - 1$$

describe the expected profit per unit bet on Outcomes 1 and 2, respectively, from the perspective of the diffuse player who believes that Outcome 1 will occur with probability p. Individual diffuse players are risk-neutral and do not affect these quantities, so they wager on an outcome only if the corresponding term is positive. For a particular diffuse player, this is true of at most one outcome as $\kappa \in (0,1)$. Consequently, if a diffuse player has identified a profitable wagering opportunity, she bets her entire fortune on it.

Despite their large space of feasible strategies, the diffuse players, aside from those whose beliefs are indexed by $p = P^{f,a}/\kappa$ or $\left(1 - P^{f,a}\right)/\kappa$, wager either 100% or 0% of their wealth on each outcome. The value of f_i at $P^{f,a}/\kappa$ and $\left(1 - P^{f,a}\right)/\kappa$ is ambiguous because, if a given diffuse player's expected profit per unit bet on Outcome i is zero, then she is indifferent to the size of her bet on Outcome i.[13]

[13] Since μ has a density, the μ-measure of a set with two points is zero. It follows that this ambiguity has no bearing on an equilibrium's uniqueness. A more serious concern is that we could have feasible strategy profiles satisfying Definition V.3 with different implied probabilities. Precluding this possibility is a key part of Theorem V.7's proof.

Though their setups differed from our own (see Sections 5.2 - 5.3), Ottaviani, Sørenson, and Watanabe found similar groupings of diffuse players in equilibrium ([182]; [234]). We return to this observation during our discussion of Corollary V.9, which roughly says that these groupings persist even when we take into account the atomic player's wagers.

Proof. There is little to formalize beyond our heuristic discussion above. We only comment that rearranging (5.7) shows that we can never have both

$$\frac{\kappa p}{P^{f,a}} - 1 > 0 \qquad \text{and} \qquad \frac{\kappa(1-p)}{1-P^{f,a}} - 1 > 0.$$

□

Proposition V.6. *Let (f,a) be a feasible strategy profile such that $d_1,\ d_2 > 0$. Consider the equation*

$$\begin{aligned} a_1 &\left(\frac{\kappa\, q}{P^{f,a}} - 1\right) + a_2\left(\frac{\kappa\,(1-q)}{1-P^{f,a}} - 1\right) \\ &= \sup_{\substack{b_1,b_2\geq 0\\ b_1+b_2\leq w}} \left\{ b_1\left(\frac{\kappa\, q}{P^{f,b}} - 1\right) + b_2\left(\frac{\kappa\,(1-q)}{1-P^{f,b}} - 1\right)\right\}. \end{aligned} \tag{5.8}$$

(i) When

$$q > \frac{d_1}{\kappa\,(d_1+d_2)}, \tag{5.9}$$

a satisfies (5.8) if and only if $a_2 = 0$ and

$$a_1 = \min\left\{w,\ \sqrt{\frac{\kappa q d_1 d_2}{1-\kappa q}} - d_1\right\}. \tag{5.10}$$

(ii) When

$$1-q > \frac{d_2}{\kappa\,(d_1+d_2)}, \tag{5.11}$$

a satisfies (5.8) if and only if $a_1 = 0$ and

$$a_2 = \min\left\{w,\ \sqrt{\frac{\kappa\,(1-q)\,d_1 d_2}{1-\kappa\,(1-q)}} - d_2\right\}. \tag{5.12}$$

(iii) When

$$q \leq \frac{d_1}{\kappa\,(d_1+d_2)} \qquad \textit{and} \qquad 1-q \leq \frac{d_2}{\kappa\,(d_1+d_2)}, \tag{5.13}$$

a satisfies (5.8) if and only if $a_1 = a_2 = 0$.

As in our discussion of the last result, $P^{f,a} \in (0,1)$ and that we only study the case in which $d_1,\ d_2 > 0$ does not matter. Proposition V.6 can be interpreted as the complement of Proposition V.5: It characterizes the atomic player's equilibrium strategy, given the diffuse players' bets.

A calculation similar to that in (5.7) shows that (5.9) and (5.11) never hold simultaneously. Notice that $a_1 > 0$ under (i), while $a_2 > 0$ under (ii). For example, (5.9) implies that

$$\frac{d_1}{d_1+d_2} < \kappa q \qquad \text{and} \qquad \frac{d_2}{d_1+d_2} > 1-\kappa q. \tag{5.14}$$

In this case,

$$a_1 = \min\left\{w,\ \sqrt{\frac{\kappa q d_1 d_2}{1-\kappa q}} - d_1\right\} \geq \min\left\{w,\ \sqrt{\left(\frac{d_1}{d_2}\right) d_1 d_2} - d_1\right\} > 0.$$

Hence, under (i), the atomic player wagers on Outcome 1 alone. The atomic player only bets on Outcome 2 in (ii), while she does not wager at all in (iii). Despite having the opportunity to do so, she never simultaneously wagers on both possibilities. Proposition V.5 revealed similar behavior for diffuse bettors.

Important ideas in the proofs of Propositions V.5 and V.6 are closely related. By rearranging (5.9) and (5.11), we get

$$\frac{\kappa\,(d_1+d_2)\,q}{d_1} - 1 > 0 \qquad \text{and} \qquad \frac{\kappa\,(d_1+d_2)\,(1-q)}{d_2} - 1 > 0,$$

respectively. Given the wagers of the diffuse players, the first term describes the expected profit per unit bet on Outcome 1 according to the atomic player. The second term has the analogous interpretation for Outcome 2.

As in our analysis for the diffuse players, the atomic player is risk-neutral and bets only when one of these inequalities holds,[14] leading directly to (iii). Isaacs first proved this while modeling parimutuel wagering as the control problem faced by a single risk-neutral atomic player unconstrained by a budget ([138]). We include (iii) in Proposition V.6 merely to assist with our presentation.

Unlike our previous analysis, we cannot conclude that the atomic player bets her entire fortune on an initially profitable wagering opportunity. The reason is simple: the atomic player's choices affect (5.2). In fact, all else being equal, the payoffs per unit bet on an outcome decrease as the atomic player raises her wager on that outcome. Balancing the desires to increase her expected profit by betting more and keep her expected profit per unit bet high by betting less leads to (5.10) and (5.12), not the all-or-nothing wagers of Proposition V.5.

More precisely, if (5.9) holds and we relax our wealth constraint, this trade-off makes it optimal for the atomic player to wager

$$\sqrt{\frac{\kappa q d_1 d_2}{1-\kappa q}} - d_1 \tag{5.15}$$

on Outcome 1. This solution was also first discovered by Isaacs ([138]).[15] The idea behind (5.10) is then clear: If the atomic player cannot afford to bet (5.15) on Outcome 1, she instead wagers as much as she can. This seems reasonable, intuitively, since up to (5.15), the positive impact of raising her Outcome 1 bet on her expected profit should outweigh the negative impact. The interpretation of (ii) is similar. We present the formal proof of Proposition V.6 in Section 5.7.

[14] The outcome, if any, on which the atomic player wagers can be identified based upon her opponents' wagers alone. In particular, this identification can be made without knowledge of the implied probability of Outcome 1. Still, from (5.8), it is clear that the atomic player bets on Outcome i in equilibrium if and only if the expected profit per unit bet on Outcome i is positive. We rigorously prove and discuss this further in Corollary V.8.

[15] Related expressions are also seen in equilibrium studies with N risk-neutral atomic players constrained to wager a specific total amount ([77]). In an obvious way, Proposition V.6 fills the small gap between these two settings. Recall that no explicit solutions are available for Cheung's atomic player, who faces a budget constraint like ours but is risk-averse ([85]).

Theorem V.7. *A pure-strategy Nash equilibrium exists if and only if* $\kappa > 0.5$.[16] *When an equilibrium exists, it is unique.*

The connection between low transaction costs (or large κ) and the existence of non-trivial equilibria has been observed in other parimutuel wagering studies. For example, a non-trivial equilibrium exists in Watanabe's model only if the house take is sufficiently small ([234]). Ottaviani and Sørenson make a similar observation ([182]).

Intuitively, $\kappa > 0.5$ is necessary for the existence of an equilibrium in our framework because there are only two possible outcomes. Suppose that $(f^\star, a^\star)$ is a pure-strategy Nash equilibrium. Regardless of her type, if a player believes that Outcome 1 will occur with probability p and wagers on Outcome 1, it should be true that her final expected profit per unit bet on Outcome 1 is positive:

$$\frac{\kappa\, p}{P^{f^\star,a^\star}} - 1 > 0.$$

Similarly, if she wagers on Outcome 2, then

$$\frac{\kappa\ (1-p)}{1 - P^{f^\star,a^\star}} - 1 > 0.$$

Positive amounts are wagered on both outcomes (see Section 5.3), implying that

$$\frac{\kappa}{P^{f^\star,a^\star}} - 1 > 0 \qquad \text{and} \qquad \frac{\kappa}{1 - P^{f^\star,a^\star}} - 1 > 0. \tag{5.16}$$

Rearranging (5.16) shows that $\kappa > 0.5$. More generally, it is easy to see that in a parimutuel betting game with n outcomes and risk-neutral players who can elect not to bet, $\kappa > 1/n$ is necessary for the existence of an equilibrium.

Parimutuel wagering games in the literature occasionally possess multiple equilibria. For instance, Watanabe et al. adapt the work of Harsanyi and Selten to select

[16] In practice, this inequality almost always holds.

one equilibrium out of several that arise in their atomic player model ([235]). As mentioned in Section 5.3, Watanabe's continuum game model can feature multiple equilibria when $\mu(\{p\}) > 0$ for some fixed p. We suspect that our equilibrium would no longer be unique, if we incorporated additional atomic players or relaxed our assumption that μ had a density, but we leave this issue to a future study.

We break our proof into 4 steps. Step 1 allows us to use Propositions V.5 and V.6. It says that in an equilibrium, the total amount wagered by the diffuse players on each outcome is always positive. Step 2 formalizes our discussion above and shows that $\kappa > 0.5$ is necessary for the existence of an equilibrium.

To finish, we find an equivalent formulation of our original problem. More precisely, after presenting some preliminary notation in Steps 3 - 4, we define our so-called *implied probability map* φ in Step 5. Our work in Steps 6 - 7 shows that this map has a unique fixed-point. Due to its construction, its fixed-point corresponds to a pure-strategy Nash equilibrium and vice versa. We can then conclude that our game has a unique equilibrium when $\kappa > 0.5$ in Steps 3 - 4.

Our approach is motivated by the following observation: an equilibrium is essentially determined by the implied probability of Outcome 1. Given this quantity, we immediately recover the diffuse players' wagers from Proposition V.5. Technically, we do not know how the diffuse players whose beliefs are indexed by $P^{f^\star,a^\star}/\kappa$ or $\left(1 - P^{f^\star,a^\star}\right)/\kappa$ behave, but this does not matter. By Proposition V.6, we then identify the atomic player's wagers.

This observation leads to the definition of φ in Step 5. In a certain sense, our recipe is only meaningful at a fixed-point of φ, which underlies the correspondence just discussed. Our complete proof of Theorem V.7 can be found in Section 5.8.

We close Section 5.4 with Corollaries V.8, V.9, and V.10. The first result says that the atomic player wagers on a particular outcome if and only if her final expected profit per unit bet on that outcome is positive. This is fairly obvious from (iii) of Definition V.3, and we basically assume this to be true during our discussion of Theorem V.7.

Still, recall that Proposition V.6 identifies the outcome, if any, on which the atomic player wagers based upon only the diffuse players' bets. For instance, according to that result, the atomic player wagers on Outcome 1 in an equilibrium $(f^\star, a^\star)$ if and only if

$$q > \frac{d_1^\star}{\kappa\left(d_1^\star + d_2^\star\right)}. \tag{5.17}$$

Since $a_1^\star > 0$, we might be concerned about the possibility that

$$\frac{P^{f^\star, a^\star}}{\kappa} \geq q > \frac{d_1^\star}{\kappa\left(d_1^\star + d_2^\star\right)}.$$

The specific form of $a_1^\star$ in (5.10) ultimately prevents this.

Notice that Proposition V.5 already *almost* implies the corresponding result for diffuse players. We say *almost* because of the undetermined behavior of the diffuse bettors whose final expected profit per unit bet on some outcome is 0. From that perspective, the atomic and diffuse players identify profitable wagering opportunities using the same criteria. Strategically, they just differ in how they size their equilibrium wagers.

Corollary V.8. *Suppose that $(f^\star, a^\star)$ is an equilibrium. Then $a_1^\star > 0$ if and only if*

$$\frac{\kappa q}{P^{f^\star, a^\star}} - 1 > 0,$$

while $a_2^\star > 0$ if and only if

$$\frac{\kappa\left(1 - q\right)}{1 - P^{f^\star, a^\star}} - 1 > 0.$$

Proof. We present the argument for the first case. The other is similar.

Recall that $d_1^\star$, $d_2^\star > 0$ (see Step 1 of Theorem V.7's proof). Assume that $a_1^\star > 0$. According to Proposition V.6, (5.9) holds and $a_2^\star = 0$. Using the notation from Step 4 of Theorem V.7's proof,

$$P^{f^\star,a^\star} \leq \frac{\zeta_1\left(P^{f^\star,a^\star}\right) + d_1^\star}{\zeta_1\left(P^{f^\star,a^\star}\right) + d_1^\star + d_2^\star}.$$

That

$$\frac{\kappa q}{P^{f^\star,a^\star}} - 1 > 0 \tag{5.18}$$

follows from Isaacs' work ([138]).

To prove the remaining direction, assume that (5.18) is satisfied. Exactly one of (5.9), (5.11), and (5.13) holds. We cannot have (5.11), since it would follow that $a_2^\star > 0$. Arguing as we just did, we would get

$$\frac{\kappa\left(1-q\right)}{1-P^{f^\star,a^\star}} - 1 > 0,$$

which would lead to

$$\frac{\kappa q}{P^{f^\star,a^\star}} - 1 < 0.$$

(5.13) cannot hold either, as this would give the contradiction

$$q \leq \frac{d_1^\star}{\kappa\left(d_1^\star + d_2^\star\right)} = \frac{P^{f^\star,a^\star}}{\kappa}.$$

Hence, (5.9) holds and $a_1^\star > 0$.

□

To explain our next result, suppose that some player is wagering on a particular outcome. If another player believes that this outcome will occur with higher probability than the original player, Corollary V.9 says that the new player also wagers on the outcome. Recall from Section 5.2 that Watanabe called an equilibrium with this property *regular* ([234]). All equilibria in Watanabe's framework and Ottaviani and Sørenson's framework are regular ([182]; [234]).

One might suspect that such a result generally holds, but this is not the case ([235]). To the best of our knowledge, regularity has only been consistently observed in the literature when each player's initial wealth is negligible ([182]; [234]). Corollary V.9 shows that our model is an example of a parimutuel wagering game in which the equilibrium is regular, even when an atomic player is active. It is an immediate consequence of our observation that both atomic and diffuse players decide to wager on some outcome by determining whether or not the final expected profit per unit bet on that outcome is positive.

Corollary V.9. *An equilibrium $(f^\star, a^\star)$ is always regular.*

Proof. Since $d_1^\star$, $d_2^\star > 0$ by Step 1 of Theorem V.7's proof, the result directly follows from Proposition V.5 and Corollary V.8.

□

Section 5.4's last result says that the implied probability of Outcome 1 tends to 0.5 as the house take approaches 0.5, regardless of the other parameters that we choose for our model. In fact, the convergence is uniform.

Initially, this finding may appear rather odd. For example, it is easy to ensure that $P^{f^\star, a^\star}$ lies between 49.9% and 50.1% when $q = 0$ and the μ-mass of $[0, 1]$ is *almost* entirely concentrated near $p = 0$. If essentially the whole population believes that Outcome 2 is guaranteed to occur, how can the total amounts wagered on each outcome be roughly equal?

Our discussion of Theorem V.7 outlines the key intuition. Simply notice that instead of rearranging (5.16) to show that $\kappa > 0.5$, we can show that $P^{f^\star, a^\star} \in$

$(1-\kappa, \kappa)$. This holds even in the extreme scenario where virtually all of the initial wealth is held by those who believe that Outcome 2 is a sure bet. Still, our first instinct has some merit: Here, $P^{f^\star, a^\star} \approx 1-\kappa$ for all κ (see Section 5.6).

Corollary V.10. *Fix μ, q, and w and consider the map defined on $(0.5, 1)$ by*

$$\kappa \mapsto P^{f^\star, a^\star}.$$

As $\kappa \downarrow 0.5$, the values of the map approach 0.5.

Proof. Simply notice that the map is well-defined by Theorem V.7 and that $P^{f^\star, a^\star} \in (1-\kappa, \kappa)$ by Step 4 of its proof.

□

5.5 Numerical Results

The theoretical results from Section 5.4 allow us to return to our central question: How do large-scale participants in parimutuel wagering events affect the house and ordinary bettors? We explore this issue by analyzing several concrete scenarios (see Examples V.12 - V.15).

We use the house's revenue to quantify the atomic player's impact on the house. Given an equilibrium $(f^\star, a^\star)$, the house's revenue is simply the product of the house take and the total amount wagered:

$$(1-\kappa)\left(d_1^\star + d_2^\star + a_1^\star + a_2^\star\right). \tag{5.19}$$

Notice that this quantity is deterministic and does not depend on the *actual* probability of Outcome 1, as the house collects (5.19) regardless of which outcome occurs.

To quantify the atomic player's effect on diffuse bettors, we use one of two quantities. In Examples V.12 - V.13, we select values for the *actual* probability of Outcome

1. Making this choice lets us compute the *actual* total expected profit of the diffuse players. If $(f^\star, a^\star)$ is an equilibrium and the *actual* probability of Outcome 1 is $\bar{p}$, then it is given by

$$d_1^\star\left(\frac{\kappa\,\bar{p}}{P^{f^\star,a^\star}} - 1\right) + d_2^\star\left(\frac{\kappa\,(1-\bar{p})}{1-P^{f^\star,a^\star}} - 1\right). \tag{5.20}$$

While we use (5.20) to describe how the atomic player affects the diffuse players in Examples V.12 - V.13, we cannot do so in Examples V.14 - V.15. The reason is that we make no assumption about the *actual* probability of Outcome 1 in the latter situations. Instead, we quantify the impact on diffuse bettors using their total *subjective* expected profit, which is given by

$$\int_0^1 \left[f_1^\star(p)\left(\frac{\kappa\,p}{P^{f^\star,a^\star}} - 1\right) + f_2^\star(p)\left(\frac{\kappa\,(1-p)}{1-P^{f^\star,a^\star}} - 1\right)\right] d\mu(p) \tag{5.21}$$

in an equilibrium $(f^\star, a^\star)$. Here, we merely compute each diffuse player's expected profit according to her beliefs and aggregate the results over all diffuse players.

From Section 6.1, recall that the standard narrative says that the presence of the atomic player should increase the house's revenue but decrease the diffuse players' total expected profit (*actual* or *subjective*, as applicable). The eventual decline in the house's revenue should only be seen over time, not in our static game model.

Technically, since we specified that $w > 0$ in Section 5.3, the atomic player can never be absent in our framework. Still, we can model the atomic player's absence by choosing an extremely low value for w, say $w = 10^{-10}$. Her wagers are then too small to materially affect any equilibrium quantities. We do this for all of the Case 1's in Examiples V.12 - V.15. The atomic player is present, that is, $w >> 0$, in all of our upcoming Case 2's.

The following collection of measures is convenient for our purposes.

Definition V.11. For $n \geq 1$, let μ_n be the Borel measure on $[0,1]$ whose density g_n is defined by

$$g_n(p) = \begin{cases} -2n(n-1)p + 2(n-1) + 1/n & \text{if } p < 1/n \\ 1/n & \text{if } p \geq 1/n \end{cases}.$$

In Figure 5.1, we give the plots of g_n for $n = 1$, 3, and 9. Here are the key observations:

(i) g_n is continuous and positive on $[0,1]$.

(ii) $\mu_n([0,1]) = 1$ for all n.

(iii) μ_1 is the Lebesgue measure on $[0,1]$.

(iv) g_n converges in distribution to the Dirac delta function as $n \uparrow \infty$.

(i) and (ii) ensure that μ_n is a suitable candidate for the measure describing the initial wealth of the diffuse bettors, that is, all of our results apply when $\mu = \mu_n$. In this case, (iii) says that the initial wealth of the diffuse bettors is uniformly distributed when $n = 1$. (iv) says that their initial wealth is increasingly concentrated among those who believe that Outcome 2 will occur with high probability as n increases. Another key feature is that the diffuse bettors' total initial wealth is always equal to 1 (see (ii)).

Before we proceed, we remark that all of our figures are generated with the help of the ideas in Theorem V.7's proof. More precisely, assume that $\kappa \in (0.5, 1)$. The function φ defined in Step 5 has a unique fixed-point. Since φ is also continuous and decreasing, we can efficiently approximate this value with arbitrary precision using binary search. Step 3 shows how to reconstruct the pure-strategy Nash equilibrium $(f^\star, a^\star)$, given such an estimate.

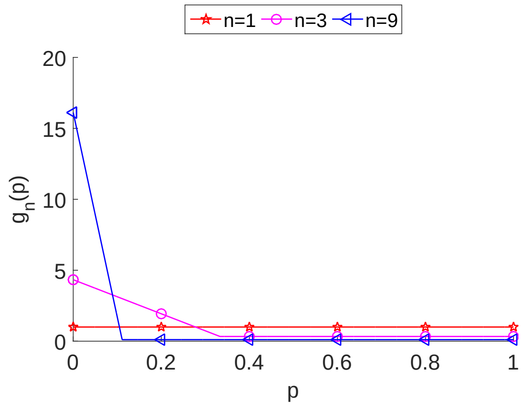

Figure 5.1: Depiction of g_n for $n = 1$, 3 and 9.

Example V.12. We now show that both effects predicted by the usual narrative can be observed.

In Cases 1 and 2, we set $\mu = \mu_1$ and $q = 0.9$. We assume that the atomic player's beliefs are exactly correct, i.e., the *actual* probability that Outcome 1 will occur is also 0.9. The only difference between Cases 1 and 2 is that $w = 10^{-10}$ in the former but $w = 1$ in the latter.

Recall that choosing $\mu = \mu_1$ means that the diffuse bettors' initial wealth is uniformly distributed. Since $q = 0.9$, the atomic player (correctly) believes that Outcome 1 is quite likely.

In Figure 5.2, we plot the diffuse players' *actual* total expected profit. Collectively, the diffuse players beliefs are rather inaccurate, so it is not surprising that their expect profit is negative. Still, as $\kappa \uparrow 1$, the diffuse players become increasingly worse off in Case 2. Intuitively, the atomic player quickly raises her bet on Outcome 1 as $\kappa \uparrow 1$. The implied probability of Outcome 1 rises,[17] causing more diffuse players to bet on Outcome 2 and less to bet on Outcome 1. Since the *actual* probability of Outcome 1 is 0.9, this transition negatively affects the diffuse players.

[17] We remark that $P^{f^\star, a^\star} \in [0.5, 0.7]$ for all $\kappa \in (0.5, 1)$ in Example V.12. In particular, the *favorite-longshot bias* results, regardless of the house take.

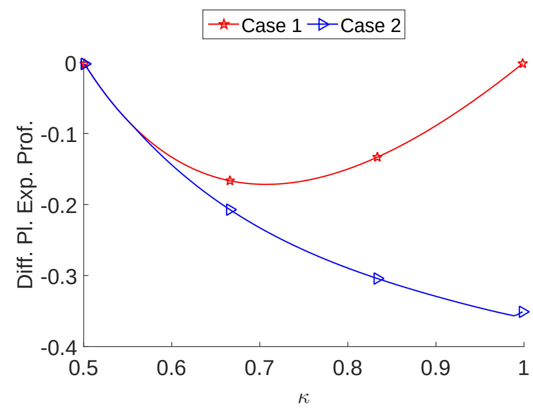

Figure 5.2: Depiction of the diffuse players' *actual* total expected profit in Example V.12.

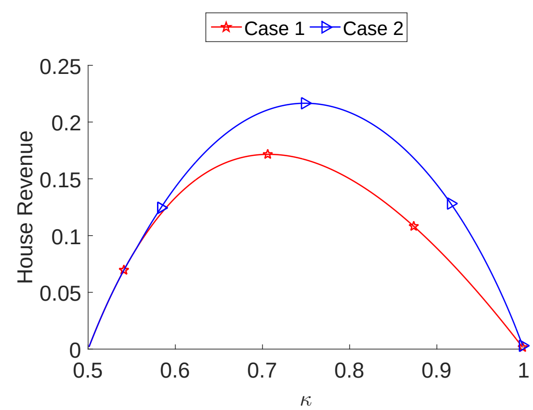

Figure 5.3: Depiction of the house's revenue in Example V.12.

We plot the house's revenue in Figure 5.3. The house's revenue is higher in Case 2 than in Case 1 for all $\kappa \in (0.5, 1)$, as a result of the higher wagering totals in Case 2.

Example V.13. Still, there are cases in which diffuse players are positively affected by the activities of the atomic player.

We now choose $\mu = \mu_1$ and $q = 0.57$ for Cases 1 and 2. The *actual* probability of Outcome 1 is 0.47. The only difference between the two scenarios is that $w = 10^{-10}$ in Case 1, while $w = 1$ in Case 2.

Compared to Example V.12, Outcome 1 is slightly less likely here. Also, we still assume that the atomic player's forecast is quite accurate, but her prediction is no

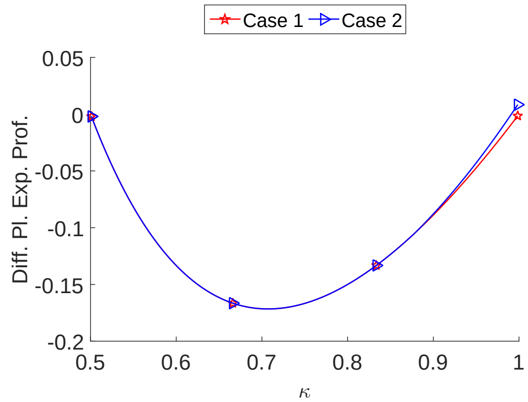

Figure 5.4: Depiction of the diffuse players' *actual* total expected profit in Example V.13.

longer perfect.

We plot the diffuse players' *actual* expected profit in Figure 5.4. The graphs for Cases 1 and 2 are the same for most values of κ; however, we see that the diffuse players' expected profit is *higher* in Case 2 when κ is large enough. Roughly, the atomic player does not start betting (on Outcome 1) until the house take is low, since she is nearly ambivalent. This explains why the curves are initially identical. When the atomic player begins to wager on Outcome 1, the diffuse players reshuffle their bets as in Example V.12.[18] The shift benefits them, essentially because the atomic player wagers on the *wrong* outcome.

In Figure 5.5, we plot the house's revenue. We see an extremely small improvement in Case 2 when κ is large, but the graphs are almost indistinguishable, visually. As in Example V.12, the increase corresponds to the increase in the total amount wagered. It is slight, as the atomic player's uncertainty about what will occur causes her to bet very little in Case 2, even for $\kappa \approx 1$.

Example V.14. We can argue that the diffuse players are better off in the presence of the atomic player, even without making assumptions about the *actual* probability of Outcome 1.

[18] Like Example V.12, we see the *favorite-longshot bias* here because $P^{f^\star, a^\star} \in [0.5, 0.53]$ for all $\kappa \in (0.5, 1)$.

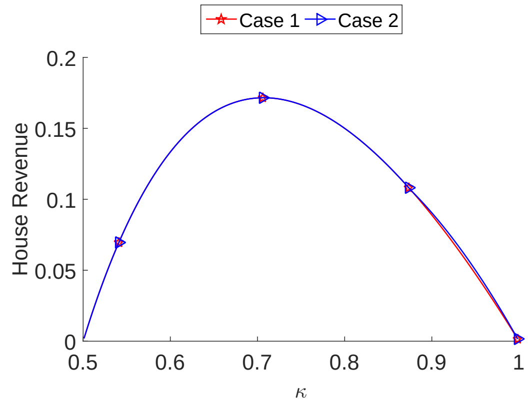

Figure 5.5: Depiction of the house's revenue in Example V.13.

In Cases 1 and 2, we now choose $\mu = \mu_{10}$ and $q = 0.95$. As in Examples V.12 - V.13, we set $w = 10^{-10}$ in Case 1 and $w = 1$ in Case 2. The atomic player believes that Outcome 1 is highly likely. Collectively, the diffuse players believe that Outcome 2 will probably happen: over 90% of their initial wealth is held by those who believe that the probability of Outcome 2 is at least 0.9. We make no judgment about the accuracy of the players' beliefs.

We plot the diffuse players' *subjective* expected profit in Figure 5.6. The graphs are the same for low κ, but eventually, the diffuse players' *subjective* expected profit is much higher in Case 2. Intuitively, the atomic player raises her wager on Outcome 1 as $\kappa \uparrow 1$, since she believes that Outcome 1 will occur. The implied probability of Outcome 1 then rises, a boon to those who believe that Outcome 2 will occur. This includes *most* of the diffuse players.

In Figure 5.7, we plot the house's revenue. The house's revenue in Case 2 is at least as large as its revenue in Case 1 for all values of κ. Often, the improvement is significant. Roughly, there is too much agreement among the diffuse bettors. They wager more in the presence of the atomic player, *believing* that they can profit from her *supposed* wagering mistakes. The pool size increases, leading to greater revenue for the house in Case 2.

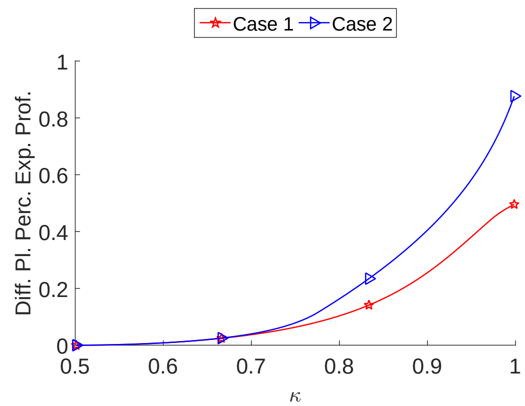

Figure 5.6: Depiction of the diffuse players' *subjective* total expected profit in Example V.14.

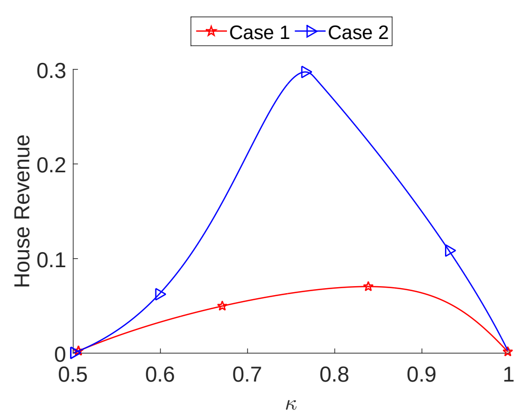

Figure 5.7: Depiction of the house's revenue in Example V.14.

Example V.15. In keeping with the standard narrative, it seems like the house is always immediately better off because of the atomic player. We can cast some doubt on this too.

We used the same μ for Cases 1 and 2 in Examples V.12 - V.14. This choice captures the idea that the presence of the atomic player should not affect the diffuse players' initial wealth in a static parimutuel wagering game.[19]

Alternatively, it might be reasonable to fix the distribution of initial wealth across the *entire* population, not just the diffuse population. For instance, we could specify that the amount held by those who believe that Outcome 1 will occur is roughly the same as the amount held by those who believe that Outcome 2 will occur. We could then compare the situations in which the wealth is held by only diffuse players and in which some wealth is held by the atomic player. This is the approach we now take.

In Case 1, $w = 10^{-10}$ and the density of μ is defined by

$$p \mapsto \frac{g_{100}\left(p\right) + g_{100}\left(1-p\right)}{2}.$$

For Case 2, we set $w = 1$ and $\mu = \mu_{100}$. We choose $q = 1$ for both scenarios and, again, make no assumption about Outcome 1's *actual* probability.

Intuitively, *half* of the diffuse players in Case 1 believe that Outcome 1 is going to occur, while the other *half* believes that Outcome 2 will occur. For Case 2, the diffuse players *all* believe that Outcome 2 is going to occur. The atomic player, whose wealth is equal to the collective wealth of the diffuse players, believes that Outcome 1 will occur.

Notice that the diffuse players' total initial wealth is equal in Cases 1 and 2; however, the wealth of the entire population in Case 2 is twice what it is in Case

[19]Over the course of many events, diffuse players may stop participating because of the atomic player, making this intuition questionable in another context (see Section 6.1).

1. This is consistent with our parameter selections in Examples V.12 - V.14. When studying the house's revenue here, we might have also chosen $\mu = 0.5 \times \mu_{100}$ and $w = 0.5$ in Case 2, ensuring that the wealth of the entire population is identical in both scenarios. We comment on this shortly.

One might suspect that Cases 1 and 2 are quite similar, but this is not true. In Figure 5.8, we plot the diffuse players' *subjective* expected profit. For $\kappa \approx 1$, their *subjective* expected profit is higher in Case 2 than in Case 1. Otherwise, it is lower for all κ, and often, the drop is significant.

To explain this observation, we plot the implied probability of Outcome 1 in Figure 5.9. The implied probability of Outcome 1 is about 0.5 for all κ in Case 1. For Case 2, it is almost $1 - \kappa$ for low κ but approaches 0.5 as $\kappa \uparrow 1$.

Since the diffuse players believe that Outcome 2 will occur but the atomic player believes that Outcome 1 will occur, the intuition appears to be as follows. Roughly, the diffuse players are willing to tolerate unfavorable betting conditions more than the atomic player. Despite the fact that each type of player is *basically* sure that the outcome they are betting on will occur, the diffuse players in Case 2 raise the size of their wagers much faster than the atomic player when κ is low. The diffuse players do this at the expense of their own *subjective* expected profit, which is nearly zero in Case 2 until $\kappa \approx 0.84$. The atomic player does not substantially raise her wager on Outcome 1 until the values of the implied probability of Outcome 1 and κ make her *subjective* expected profit per unit bet very high.

One could argue that the atomic player's strategy is superior to the strategies employed by the diffuse players, although we have made no assumption about the accuracy of her prediction. Perhaps the reason is that unlike the diffuse players, she considers her individual impact on the other wagerers because of her substantial wealth.

In Figure 5.10, we plot the house's revenue. We see that the house's revenue is higher in Case 1 for small κ but higher in Case 2 for large κ. After selecting $\mu = 0.5 \times \mu_{100}$ and $w = 0.5$ in Case 2 (see our explanation above), we re-plot the house's revenue in Figure 5.11. Now the house's revenue is higher in Case 1 for all κ. Regardless of our normalization, the key point is that the house's maximum revenue is always much higher in Case 1.

Intuitively, the diffuse players appear to have a greater tolerance for poor betting conditions than the atomic player, as discussed previously. They are willing to bet even when $\kappa \approx 0.5$, and consequently, about 99% of the diffuse players are wagering in Case 1 when $\kappa \approx 0.51$. The house loses revenue by raising κ, since the entire population has already wagered *almost* everything that it can. In Case 2, the atomic player's reluctance to significantly raise her wager on Outcome 1 until betting conditions improve means that the total amount wagered is low for most κ. The pool is large only if κ is quite high, so the house does not collect much.

In Examples V.12 - V.14, the house's revenue in Case 2 is at least as great as the house's revenue in Case 1 for all κ. Of course, the house's maximum revenue is then higher in Case 2 for these studies. Here, the pointwise analysis is not as clean, leading us to more directly consider the house's strategic behavior.

Thus far, we have understood the house take to be exogenously determined. One way to relax this assumption is to break our game into two stages. In the first stage, the betting organizer chooses a value for the house take in order to maximize her revenue. The second stage is identical to our current setup.

Under the new framework, the house needs to account for more than the distribution of initial wealth across the entire population. The house must also consider

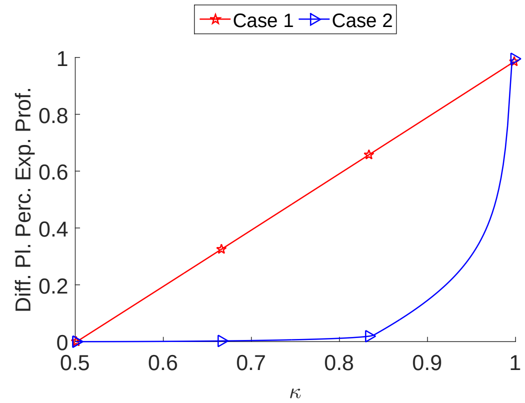

Figure 5.8: Depiction of the diffuse players' *subjective* total expected profit in Example V.15.

how the initial wealth is distributed across the population *for each type of player*. Independent of our normalization of the entire population's wealth, the house's revenue is maximized at $\kappa \approx 0.506$ in Case 1 and at $\kappa \approx 0.839$ in Case 2. Hence, the house selects these values of κ in Cases 1 and 2, respectively.

In Figure 5.8, the diffuse players' *subjective* expected profit is about 0.0085 when $\kappa \approx 0.506$ in Case 2, while it is about 0.023 when $\kappa \approx 0.839$ in Case 2. These numbers arise from our original parameter choices, ensuring that the total initial wealth of the diffuse players is 1 in both scenarios. The diffuse players are not well off in either Case 1 or Case 2. We could still argue that the setup of Case 2 is to their advantage, in contrast to our earlier pointwise analysis. Roughly, because the atomic player is less willing to bet under poor wagering conditions, the house more easily preys upon the diffuse players when she is absent.

5.6 Appendix: Properties of $P^{f^\star, a^\star}$

We briefly revisit our discussion of Corollary V.10 by studying the properties of $P^{f^\star, a^\star}$ as a function of κ. We hope this highlights a few theoretical features of our model, though we do not relate our findings in Section 5.6 to our central question. To generate these figures, we use the approach and notation of Section 5.5.

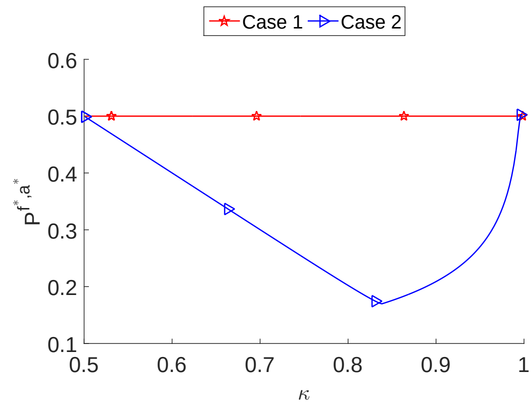

Figure 5.9: Depiction of Outcome 1's implied probability in Example V.15.

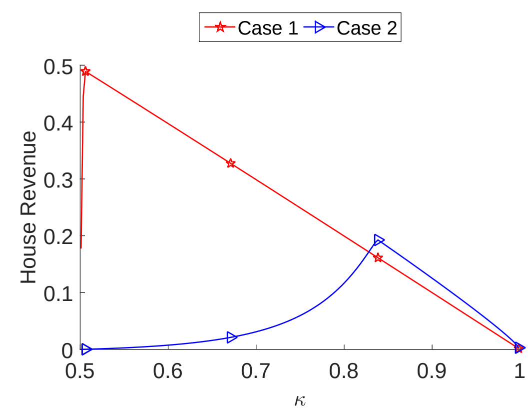

Figure 5.10: Depiction of the house's revenue in Example V.15 (before Case 2 wealth is normalized).

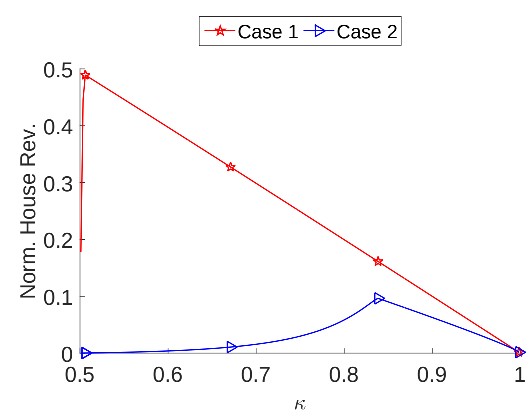

Figure 5.11: Depiction of the house's revenue in Example V.15 (after Case 2 wealth is normalized).

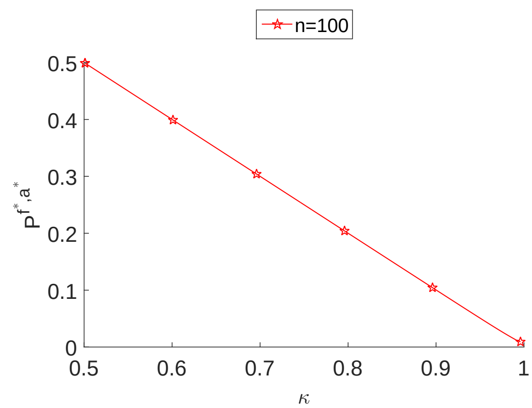

Figure 5.12: Depiction of Outcome 1's implied probability when all players effectively believe that Outcome 2 will occur (cf. Section 5.6).

Recall our claim that in an extreme case where nearly all of the initial wealth is held by those who believe that Outcome 2 is highly likely, $P^{f^\star,a^\star} \approx 1 - \kappa$ for all κ. We illustrate this in Figure 5.12, which depicts the map $\kappa \mapsto P^{f^\star,a^\star}$ when $\mu = \mu_{100}$, $q = 0$, and $w = 1$. Here, over 99% of the diffuse players' initial wealth belongs to those who believe that the probability of Outcome 2 is at least 0.99. The atomic player believes that Outcome 2 is guaranteed to occur. As expected, the plot is visually indistinguishable from a plot of the map $\kappa \mapsto 1 - \kappa$.

Often, it is more difficult to broadly describe attributes of the function $\kappa \mapsto P^{f^\star,a^\star}$. We know that $P^{f^\star,a^\star} \in (1-\kappa, \kappa)$, but our players' heterogeneity allows for a wide range of possibilities within these bounds. There is a rich interplay between their differing beliefs, effects on (5.2), and wealth constraints. By plotting the map $\kappa \mapsto P^{f^\star,a^\star}$ under varying assumptions on q, w, and μ, Figure 5.13 displays a few of the myriad possibilities. For instance, the density g of the μ used to generate Line A is a linear combination of Gaussian densities with different means. Line A's oscillations arise because of g's distinct peaks.

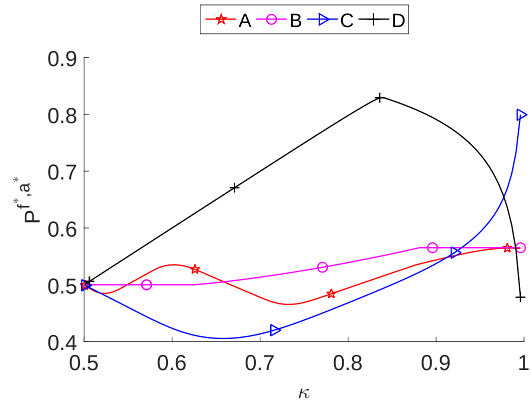

Figure 5.13: Depiction of Outcome 1's implied probability under various parameter regimes (cf. Section 5.6).

5.7 Appendix: Proof of Proposition V.6

We only prove (i). The argument for (ii) is similar, and as observed above, (iii) is due to Isaacs ([138]).

Suppose that (5.9) holds. Then

$$\frac{1-q}{d_2} < \frac{1-\kappa q}{d_2} < \frac{1}{d_1+d_2} < \frac{1}{\kappa\left(d_1+d_2\right)} < \frac{q}{d_1}. \tag{5.22}$$

We get the first and third inequalities because $\kappa \in (0,1)$. The second inequality is due to (5.14), while the last inequality is a rearrangement of (5.9). For now, the critical observation is that the leftmost quantity is less than the rightmost, allowing us to use the work of Isaacs ([138]).

Define the map

$$\Phi : \mathbb{R}^2_{\geq 0} \longrightarrow \mathbb{R}$$

by

$$\Phi\left(b_1, b_2\right) = b_1\left(\frac{\kappa\left(b_1+d_1+b_2+d_2\right)q}{b_1+d_1} - 1\right) + b_2\left(\frac{\kappa\left(b_1+d_1+b_2+d_2\right)\left(1-q\right)}{b_2+d_2} - 1\right).$$

$\Phi\left(b_1, b_2\right)$ is the atomic player's expected profit, given that she wagers b_i on Outcome

i, and it allows us to rewrite (5.8) as

$$a_1\left(\frac{\kappa\, q}{P^{f,a}} - 1\right) + a_2\left(\frac{\kappa\ (1-q)}{1-P^{f,a}} - 1\right) = \sup_{\substack{b_1,b_2\geq 0\\ b_1+b_2\leq w}} \Phi\left(b_1, b_2\right).$$

Technically, strategy profiles (b_1, b_2) with

$$b_1 + b_2 > w$$

are not feasible, but this is unimportant.

Isaacs showed that Φ has a unique global maximum on $\mathbb{R}^2_{\geq 0}$, denoted $(b_1^\star, b_2^\star)$, given by

$$(b_1^\star, b_2^\star) = \left(\sqrt{\frac{\kappa q d_1 d_2}{1-\kappa q}} - d_1, 0\right).$$

This proves that (i) holds when $w \geq b_1^\star$.

The case where $w < b_1^\star$ is handled with elementary calculus. First, observe that (5.9) implies that $q > 0$ and

$$\partial_{b_1}\Phi\left(0,0\right) = \frac{\kappa\left(d_1+d_2\right)q}{d_1} - 1 > 0.$$

Since $\partial_{b_1}\Phi\left(b_1^\star, 0\right) = 0$ and

$$\partial_{b_1 b_1}\Phi\left(b_1, b_2\right) = -\frac{2\kappa d_1\left(b_2+d_2\right)q}{\left(b_1+d_1\right)^3}$$

is always negative, it follows that $\partial_{b_1}\Phi\left(b_1, 0\right) > 0$ for $0 \leq b_1 < b_1^\star$. The interpretation is that as long as the atomic player has not yet wagered on Outcome 2, she should wager as much as she can up to $b_1^\star$ on Outcome 1.

We only need to argue that she should never wager on Outcome 2, that is, $\partial_{b_2}\left(b_1, b_2\right)$ is negative whenever $b_1 + b_2 \leq w$. Rearranging (5.22) implies that

$$\partial_{b_2}\Phi\left(0,0\right) = \frac{\kappa\left(d_1+d_2\right)\left(1-q\right)}{d_2} - 1 < 0.$$

Now $\partial_{b_2}\Phi\left(b_1^\star, 0\right) \leq 0$[20] and

$$\partial_{b_2 b_1}\Phi\left(b_1, b_2\right) = \frac{\kappa d_1 q}{\left(b_1 + d_1\right)^2} + \frac{\kappa d_2\left(1-q\right)}{\left(b_2 + d_2\right)^2}$$

is positive everywhere, so $\partial_{b_2}\Phi\left(b_1, 0\right) < 0$ for $0 \leq b_1 < b_1^\star$. Since

$$\partial_{b_2 b_2}\Phi\left(b_1, b_2\right) = -\frac{2\kappa d_2\left(b_1 + d_1\right)\left(1-q\right)}{\left(b_2 + d_2\right)^3}$$

is always non-positive, we conclude that $\partial_{b_2}\left(b_1, b_2\right)$ must be negative whenever $0 \leq b_1 < b_1^\star$. In particular, $\partial_{b_2}\left(b_1, b_2\right)$ is negative, if $b_1 + b_2 \leq w$.

□

5.8 Appendix: Proof of Theorem V.7

Step 1: If $(f^\star, a^\star)$ is an equilibrium, then $d_1^\star$, $d_2^\star > 0$.

We show that the total amount wagered on each outcome by the diffuse players is positive in equilibrium: $d_1^\star$, $d_2^\star > 0$. We use this result to complete the proof of the *only if* direction of Theorem V.7 in Step 2. It also allows us to use Propositions V.5 and V.6 in the proof of the *if* direction.

Suppose instead that we can find a pure-strategy Nash equilibrium $(f^\star, a^\star)$ with $d_1^\star = 0$. It follows immediately that $d_2^\star > 0$: otherwise, (iii) of Definition V.3 implies that $a_1^\star = a_2^\star = 0$, which contradicts (i) of Definition V.3. A consequence of Isaacs' work is that $q = 0$ ([138]).[21] Then (iii) of Definition V.3 implies that $a_1^\star = a_2^\star = 0$. In particular, $P^{f^\star, a^\star} = 0$. By (i) of Definition V.3, $f_2^\star \equiv 0$, which is impossible since $d_2^\star > 0$.

Hence, $d_1^\star > 0$. It follows similarly that $d_2^\star > 0$.

Step 2: If an equilibrium exists, then $\kappa > 0.5$.

[20] We know that $\partial_{b_1}\Phi\left(b_1^\star, 0\right) = 0$ and $\partial_{b_2}\Phi\left(b_1^\star, 0\right) \leq 0$ since Φ has a unique global maximum on $\mathbb{R}^2_{\geq 0}$ at $\left(b_1^\star, b_2^\star\right)$.

[21] If $q > 0$, then $a_1^\star$ is undefined. Roughly, the atomic player needs to make an arbitrarily small bet on Outcome 1 but is not allowed to do so.

We finish the proof of the *only if* direction of Theorem V.7 by formalizing the heuristics given previously. Suppose that $(f^\star, a^\star)$ is an equilibrium. Step 1 implies that $d_1^\star$, $d_2^\star > 0$. Then both

$$\frac{P^{f^\star,a^\star}}{\kappa} < 1 \qquad \text{and} \qquad \frac{1 - P^{f^\star,a^\star}}{\kappa} < 1 \tag{5.23}$$

by Proposition V.5. Rearranging (5.23) finishes the argument.

Step 3: Definition and discussion of $\bar{p}_i$ (when $\kappa > 0.5$).

Due to Step 2, we assume that $\kappa > 0.5$ for the remainder of the proof (Steps 3 - 4). This assumption ensures that our discussions are not vacuous, as we implicitly rely on the positive length of the interval $[1-\kappa, \kappa]$.

We now define and discuss the quantities $\bar{p}_1$, $\bar{p}_2 \in [1-\kappa, \kappa]$. We use this notation when we define our ζ_i maps in Step 4 and φ in Step 5. Roughly, $\bar{p}_1$ is a naïve lower bound for the implied probability of Outcome 1 when the atomic player wagers on Outcome 1. Similarly, $\bar{p}_2$ is an upper bound for the implied probability of Outcome 1 when the atomic player wagers on Outcome 2. Both are derived from Propositions V.5 and V.6.

Since μ's density is positive, the map

$$p \mapsto \frac{\mu\left(\frac{p}{\kappa}, 1\right]}{\kappa\left(\mu\left(\frac{p}{\kappa}, 1\right] + \mu\left[0, 1 - \frac{1-p}{\kappa}\right)\right)} \tag{5.24}$$

is decreasing and continuous on $[1-\kappa, \kappa]$. Its value is $1/\kappa$ at $p = (1-\kappa)$ and 0 at $p = \kappa$. Hence, there is a unique $\bar{p}_1 \in (1-\kappa, \kappa]$ such that

$$q = \frac{\mu\left(\frac{\bar{p}_1}{\kappa}, 1\right]}{\kappa\left(\mu\left(\frac{\bar{p}_1}{\kappa}, 1\right] + \mu\left[0, 1 - \frac{1-\bar{p}_1}{\kappa}\right)\right)}. \tag{5.25}$$

Clearly, $\bar{p}_1 = \kappa$, or equivalently, $(\bar{p}_1, \kappa]$ is empty, if and only if $q = 0$. Regardless of q's value, we know

$$\begin{cases} q \leq \dfrac{\mu\left(\frac{p}{\kappa},1\right]}{\kappa\left(\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)\right)} & \text{if} \quad p\in[1-\kappa,\bar{p}_1] \\ \\ q > \dfrac{\mu\left(\frac{p}{\kappa},1\right]}{\kappa\left(\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)\right)} & \text{if} \quad p\in(\bar{p}_1,\kappa] \end{cases} . \tag{5.26}$$

We later connect this observation to (5.9) and (5.13) when defining ζ_i and φ.

Similarly, the map

$$p \mapsto \frac{\mu\left[0,1-\frac{1-p}{\kappa}\right)}{\kappa\left(\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)\right)} \tag{5.27}$$

is increasing and continuous on $[1-\kappa,\kappa]$. Its value is 0 at $p=(1-\kappa)$ and $1/\kappa$ at $p=\kappa$, so there is a unique $\bar{p}_2\in[1-\kappa,\kappa)$ such that

$$1-q=\frac{\mu\left[0,1-\frac{1-\bar{p}_2}{\kappa}\right)}{\kappa\left(\mu\left(\frac{\bar{p}_2}{\kappa},1\right]+\mu\left[0,1-\frac{1-\bar{p}_2}{\kappa}\right)\right)}.$$

Now $\bar{p}_2=(1-\kappa)$, i.e., $[1-\kappa,\bar{p}_2)$ is empty, if and only if $q=1$. In any case,

$$\begin{cases} 1-q > \dfrac{\mu\left[0,1-\frac{1-p}{\kappa}\right)}{\kappa\left(\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)\right)} & \text{if} \quad p\in[1-\kappa,\bar{p}_2) \\ \\ 1-q \leq \dfrac{\mu\left[0,1-\frac{1-p}{\kappa}\right)}{\kappa\left(\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)\right)} & \text{if} \quad p\in[\bar{p}_2,\kappa] \end{cases} . \tag{5.28}$$

This comment relates to (5.11) and (5.13), as reflected in our definitions of ζ_i and φ.

We conclude by observing that we have $\bar{p}_2<\bar{p}_1$ since $\kappa\in(0.5,1)$. This is another key remark for our future definition of φ.

Step 4: Definition and discussion of ζ_i (when $\kappa>0.5$).

We define and discuss the functions ζ_1 and ζ_2. We use this notation in our definition of φ in Step 5. Intuitively, $\zeta_i(p)$ represents the amount that the atomic player wagers on Outcome i when she makes a positive wager on Outcome i, does not face a budget constraint, and the implied probability of Outcome 1 is p (cf. (5.10) and (5.12)).

For $p \in [\bar{p}_1, \kappa]$, define ζ_1 by

$$\zeta_1(p) = \sqrt{\frac{\kappa q}{1-\kappa q}\,\mu\left(\frac{p}{\kappa},1\right]\mu\left[0,1-\frac{1-p}{\kappa}\right)} - \mu\left(\frac{p}{\kappa},1\right].$$

Since μ has a positive density, (5.26) implies that

$$\mu\left(\frac{p}{\kappa},1\right] > 0, \qquad \mu\left[0,1-\frac{1-p}{\kappa}\right) > 0, \qquad \text{and} \qquad q > \frac{\mu\left(\frac{p}{\kappa},1\right]}{\kappa\left(\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)\right)}$$

for $p \in (\bar{p}_1, \kappa)$. It follows as in (5.14) that ζ_1 is positive on $(\bar{p}_1, \kappa)$. We also have

$$\zeta_1(\bar{p}_1) = \zeta_1(\kappa) = 0$$

from (5.25).

We define ζ_2 for $p \in [1-\kappa, \bar{p}_2]$ by

$$\zeta_2(p) = \sqrt{\frac{\kappa(1-q)}{1-\kappa(1-q)}\,\mu\left(\frac{p}{\kappa},1\right]\mu\left[0,1-\frac{1-p}{\kappa}\right)} - \mu\left[0,1-\frac{1-p}{\kappa}\right).$$

Using the techniques from our discussion of ζ_1, we see that ζ_2 is positive on $(1-\kappa, \bar{p}_2)$ and

$$\zeta_2(1-\kappa) = \zeta_2(\bar{p}_2) = 0.$$

Step 5: Definition and discussion of φ (when $\kappa > 0.5$).

We introduce the *implied probability map* φ. We see in Steps 3 and 4 that a fixed-point of φ corresponds to a pure-strategy Nash equilibrium in an obvious way and vice versa, which ultimately allows us to complete Theorem V.7's proof.

φ's domain is the set of candidates p for the implied probability of Outcome 1. We need only consider $p \in [1-\kappa, \kappa]$, as we observe in Step 4. Proposition V.5 says that

$$\mu\left(\frac{p}{\kappa},1\right] \qquad \text{and} \qquad \mu\left[0,1-\frac{1-p}{\kappa}\right) \tag{5.29}$$

are the total amounts wagered by the diffuse players on Outcomes 1 and 2, respectively. The atomic player's wagers are described by Proposition V.6. For example, if

$$q > \frac{\mu\left(\frac{p}{\kappa},1\right]}{\kappa\left(\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)\right)},$$

that is, $p \in (\bar{p}_1, \kappa]$, then the atomic player wagers nothing on Outcome 2 and

$$\min\{w, \zeta_1(p)\}$$

on Outcome 1. Recalculating the implied probability of Outcome 1 using these bets, we get

$$\frac{\min\{w,\zeta_1(p)\}+\mu\left(\frac{p}{\kappa},1\right]}{\min\{w,\zeta_1(p)\}+\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)}$$

from Definition V.2. We set $\varphi(p)$ to this value. In some sense, this procedure is only potentially meaningful when p is equal to $\varphi(p)$, leading to our focus on fixed-points.

Here is the complete definition of φ suggested by this explanation:

$$\varphi(p) = \begin{cases} \dfrac{\mu\left(\frac{p}{\kappa},1\right]}{\min\{w,\zeta_2(p)\}+\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)} & \text{if} \quad p \in [1-\kappa, \bar{p}_2) \\ \dfrac{\mu\left(\frac{p}{\kappa},1\right]}{\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)} & \text{if} \quad p \in [\bar{p}_2, \bar{p}_1] \\ \dfrac{\min\{w,\zeta_1(p)\}+\mu\left(\frac{p}{\kappa},1\right]}{\min\{w,\zeta_1(p)\}+\mu\left(\frac{p}{\kappa},1\right]+\mu\left[0,1-\frac{1-p}{\kappa}\right)} & \text{if} \quad p \in (\bar{p}_1, \kappa] \end{cases}.$$

Observe that φ is continuous on $[1-\kappa, \kappa]$ since μ has a positive density and $\zeta_i(\bar{p}_i) = 0$ (see Step 4). This helps us prove that φ has a unique fixed-point in Step 7.

Step 6: φ is decreasing (when $\kappa > 0.5$).

We show that φ is decreasing. We use this property in Step 7 to argue that φ has a unique fixed-point.

Since μ's density is positive everywhere, φ is decreasing on $[\bar{p}_2, \bar{p}_1]$. It suffices to show that φ is decreasing on both $[1-\kappa, \bar{p}_2)$ and $(\bar{p}_1, \kappa]$ because φ is continuous. The proofs are similar, so we consider the former case.

We are done, if $[1-\kappa, \bar{p}_2)$ is empty. Suppose that it is not. Recall from Step 3 that this is equivalent to assuming that $q < 1$. Define two functions φ_2 and φ_2^w on $[1-\kappa, \bar{p}_2)$ by

$$\varphi_2(p) = \frac{\mu\left(\frac{p}{\kappa}, 1\right]}{\zeta_2(p) + \mu\left(\frac{p}{\kappa}, 1\right] + \mu\left[0, 1 - \frac{1-p}{\kappa}\right)}$$
$$\varphi_2^w(p) = \frac{\mu\left(\frac{p}{\kappa}, 1\right]}{w + \mu\left(\frac{p}{\kappa}, 1\right] + \mu\left[0, 1 - \frac{1-p}{\kappa}\right)}.$$

The point is that

$$\varphi(p) = \max\left\{\varphi_2(p), \varphi_2^w(p)\right\}$$

on $[1-\kappa, \bar{p}_2)$, so it is enough to show that φ_2 and φ_2^w are both decreasing.

Clearly, φ_2^w is decreasing. Denoting the positive and continuous density of μ by g, we see that φ_2 is decreasing because

$$\begin{aligned}\varphi_2'(p) = &-\frac{g\left(\frac{p}{\kappa}\right)\sqrt{\frac{\kappa(1-q)}{1-\kappa(1-q)}\mu\left(\frac{p}{\kappa}, 1\right]\mu\left[0, 1-\frac{1-p}{\kappa}\right)}}{2\kappa\left(\sqrt{\frac{\kappa(1-q)}{1-\kappa(1-q)}\mu\left(\frac{p}{\kappa}, 1\right]\mu\left[0, 1-\frac{1-p}{\kappa}\right)} + \mu\left(\frac{p}{\kappa}, 1\right]\right)^2} \\ &-\frac{\left(\frac{\kappa(1-q)}{1-\kappa(1-q)}\right)^{0.5}\mu\left(\frac{p}{\kappa}, 1\right]^{3/2}\mu\left[0, 1-\frac{1-p}{\kappa}\right)^{-0.5} g\left(1-\frac{1-p}{\kappa}\right)}{2\kappa\left(\sqrt{\frac{\kappa(1-q)}{1-\kappa(1-q)}\mu\left(\frac{p}{\kappa}, 1\right]\mu\left[0, 1-\frac{1-p}{\kappa}\right)} + \mu\left(\frac{p}{\kappa}, 1\right]\right)^2}.\end{aligned}$$

Step 7: φ has a unique fixed-point (when $\kappa > 0.5$).

We show that φ has a unique fixed-point. We use the existence of φ's fixed-point to prove the existence of a pure-strategy Nash equilibrium in Step 3, while we use the uniqueness of φ's fixed-point to demonstrate that the equilibrium is unique in Step 4.

The proof is simple: $\varphi(1-\kappa) = 1$ and $\varphi(\kappa) = 0$. Since φ is continuous and decreasing (see Steps 5 - 6), it must have a unique fixed-point.

Step 8: An equilibrium exists (when $\kappa > 0.5$).

We show that a pure-strategy Nash equilibrium exists. Based on Step 7, we need only describe how to construct an equilibrium from a fixed-point of φ.

Suppose that $\hat{P}$ is a fixed-point of φ. The proofs for each case are similar, so we only present the argument when $\hat{P} \in [1-\kappa, \bar{p}_2)$. Define a feasible strategy profile (f, a) by

$$f_1(p) = \begin{cases} 1 & \text{if } \; p \geq \hat{P}/\kappa \\ \\ 0 & \text{if } \; p < \hat{P}/\kappa \end{cases}, \qquad f_2(p) = \begin{cases} 1 & \text{if } \; 1-p \geq \left(1-\hat{P}\right)/\kappa \\ \\ 0 & \text{if } \; 1-p < \left(1-\hat{P}\right)/\kappa \end{cases},$$

$a_1 = 0$, and

$$a_2 = \min\left\{w, \zeta_2\left(\hat{P}\right)\right\}.$$

In particular,

$$d_1 = \mu\left(\frac{\hat{P}}{\kappa}, 1\right] \qquad \text{and} \qquad d_2 = \mu\left[0, 1 - \frac{1-\hat{P}}{\kappa}\right).$$

(5.28) implies that (5.11) is satisfied. Since $\varphi(1-\kappa) = 1$, we know that $\hat{P} \neq (1-\kappa)$. Hence, d_1, $d_2 > 0$ because the density of μ is positive everywhere. We then have (iii) of Definition V.3 by Proposition V.6.

Since $\hat{P}$ is a fixed-point of φ,

$$\hat{P} = \frac{\mu\left(\frac{\hat{P}}{\kappa}, 1\right]}{\min\left\{w, \zeta_2\left(\hat{P}\right)\right\} + \mu\left(\frac{\hat{P}}{\kappa}, 1\right] + \mu\left[0, 1 - \frac{1-\hat{P}}{\kappa}\right)} = P^{f,a}.$$

From Proposition V.5, we see that (ii) of Definition V.3 holds.

This completes the proof, as (i) of Definition V.3 is obviously satisfied.

Step 9: The equilibrium in Step 3 is unique (when $\kappa > 0.5$).

We conclude Theorem V.7's proof by showing that the equilibrium in Step 3 is unique. The key observation is that φ's fixed-point is also unique (see Step 7).

First, we describe how to construct a fixed-point of φ, given an equilibrium $(f^\star, a^\star)$. From Proposition V.5 and Step 1, we know that

$$d_1^\star = \mu\left(\frac{P^{f^\star,a^\star}}{\kappa}, 1\right] > 0 \qquad \text{and} \qquad d_2^\star = \mu\left[0, 1 - \frac{1 - P^{f^\star,a^\star}}{\kappa}\right) > 0.$$

In particular, $P^{f^\star,a^\star} \in (1-\kappa, \kappa)$. By Proposition V.6, there are three possibilities for $a^\star$.

Assume that $a_1^\star > 0$ and $a_2^\star = 0$. The other cases can be handled similarly. By Proposition V.6, (5.9) holds. Hence, $P^{f^\star,a^\star} \in (\bar{p}_1, \kappa)$ due to (5.26) and

$$a_1^\star = \min\left\{w, \sqrt{\frac{\kappa q d_1^\star d_2^\star}{1-\kappa q}} - d_1^\star\right\} = \min\left\{w, \zeta_1\left(P^{f^\star,a^\star}\right)\right\}.$$

By Definition V.2,

$$\begin{aligned} P^{f^\star,a^\star} &= \frac{\min\left\{w, \zeta_1\left(P^{f^\star,a^\star}\right)\right\} + \mu\left(\frac{P^{f^\star,a^\star}}{\kappa}, 1\right]}{\min\left\{w, \zeta_1\left(P^{f^\star,a^\star}\right)\right\} + \mu\left(\frac{P^{f^\star,a^\star}}{\kappa}, 1\right] + \mu\left[0, 1 - \frac{1-P^{f^\star,a^\star}}{\kappa}\right)} \\ &= \varphi\left(P^{f^\star,a^\star}\right), \end{aligned}$$

that is, $P^{f^\star,a^\star}$ is a fixed-point of φ.

Now suppose that we have another equilibrium $(f^\diamond, a^\diamond)$. Using the method just described, we see that $P^{f^\diamond,a^\diamond}$ is a fixed-point of φ. Since there is exactly one fixed-point of φ by Step 7, $P^{f^\diamond,a^\diamond} = P^{f^\star,a^\star}$.

By Step 1 and Proposition V.5, $f^\star$ and $f^\diamond$ necessarily agree everywhere, except when

$$p = P^{f^\star,a^\star}/\kappa \qquad \text{or} \qquad 1 - p = \left(1 - P^{f^\star,a^\star}\right)/\kappa.$$

Since μ has a density, it follows that $f^\star = f^\diamond$ μ-a.s. Clearly, $d_1^\star = d_1^\diamond$ and $d_2^\star = d_2^\diamond$. Proposition V.6 implies that $a^\star = a^\diamond$.

$\square$

CHAPTER VI

Mini-Flash Crashes, Model Risk, and Optimal Execution

6.1 Overview

Amidst the violent market disruption on May 6, 2010, the infamous Flash Crash,

> "Over 20,000 trades across more than 300 securities were executed at prices more than 60% away from their values just moments before. Moreover, many of these *trades were executed at prices of a penny or less, or as high as $100,000,* before prices of those securities returned to their 'pre-crash' levels" ([3]).

Today, this particular event remains so memorable due to its remarkable scale.

In fact, lesser versions of the Flash Crash, or *mini-flash crashes*, happen quite often. Anecdotal evidence suggests that there may be over a dozen every day ([102]). A rigorous empirical analysis uncovered "18,520 crashes and spikes with durations less than 1,500 ms" in stock prices from 2006 through 2011 ([142]). The exhaustive documentation on Nanex LLC's "NxResearch" site offers further corroboration as well ([10]).

A popular definition characterizes a mini-flash crash as an event in which the price of some security changes at least 0.8% and ticks ten times consecutively in a single direction ([142]). Price swings need not be so mild, though. Johnson et al. noted that "both crashes and spikes are typically more than 30 standard deviations larger than the average price movement either side of an event" ([142]). The SEC also recently described several 99% plunges as "mini-flash crashes" ([17]). Price surges can also fit these requirements. For example, the share price of Kraft Foods underwent a mini-flash crash on October 3, 2012 when it rocketed up 28% in less than a minute ([199]).

Now, it is true that the most irregular trades executed during some mini-flash crashes are eventually nullified and removed from the consolidated tape. For instance, when shares of the network security firm Qualys, Inc. jumped from \$10 to \$0.0001 and back during a 300ms period on April 25, 2013, all trades below \$10.15 were ultimately canceled ([9]). The idea is that these transactions were *clearly erroneous*, that is, there was "an obvious error in [a] term, such as price, number of shares or other unit of trading, or identification of the security" ([13]).

Regardless of whether they are reflected in final data feeds, why do such phenomena occur?

Several answers have been proposed. Roughly, most point to human errors, endogenous feedback loops, the nature of modern liquidity provision, fundamental value shocks, or market fragmentation. These ideas can be viewed as different ways to rationalize how an extreme (local or global) dislocation in supply and demand can arise in modern markets. We will thoroughly review them all in Subsection 6.2.1.

One of our contributions in the present paper is the development of a model which captures aspects of the first three theories. The remaining explanations are plausible sources of a subset of mini-flash crashes, and we discuss their relationship to our framework in Subsection 6.2.1.

Our model also appears to exhibit features of historical mini-flash crashes. For instance, there are periods in which extreme price moves will not manifest. If they do, accompanying trade volumes can be high or low. Some market participants may partially synchronize their trading during a mini-flash crash. Our agents may not know that a mini-flash crash is about to begin even just before its onset.

Our results seem to be aligned with intuitive expectations as well. For example, our mini-flash crashes can begin if some of our agents are too uncertain about their initial beliefs, inaccurate in their understanding of price dynamics, slow to update their models and objectives, or willing to take on risk.

Subsection 6.2.1 contains details on where to find the proofs and figures corresponding to these claims.

We construct our model beginning with a finite population of agents trading in a single risky asset, each of whom must decide how to act based upon his own preferences, beliefs, and observations. Our specifications are drawn from ideas in the price impact and optimal execution literature and are given in Subsections 6.4.2 - 6.4.4.

We imagine that our agents' orders are submitted to a single venue, where they are executed together with trades from other (unmodeled) market participants. This naturally compels us to make an explicit distinction between how the risky asset's price actually evolves and our agents' beliefs about its future evolution (see Section 6.5).

Since we view our agents as simultaneously solving their own optimal execution problems, we avoid certain strong assumptions that would have been implicitly needed, if we had used a classical equilibrium-based approach instead. An additional consequence is that we precisely describe the errors in our agents' beliefs. Potentially, each agent could be wrong both about how his trades affect prices and how prices would move in his absence. By appealing to theoretical and practical considerations, we argue away the consistency issues that one might feel would arise.

To the best of our knowledge, this general setup appears to be a new paradigm for modeling heterogeneous agent systems in the contexts of optimal execution and mini-flash crashes.

Additionally, we feel that our general framework could be viewed as a novel method for understanding, to some extent, model misspecification risks and Knightian uncertainty in optimal trading. The basic point is that "all models are wrong" and some (most) risks may be "unquantifiable" ([59], [150]). Existing techniques for managing these unknowns often involve position limits, sensitivity analysis, Bayesian model averaging, the worst-case framework, and interpolations between the worst-case and classical setups. We illustrate how employing our abstract process, potentially in conjunction with these standard methods, may give a more robust perspective.

Our in-depth discussion of these contributions and connections to previous literature on optimal execution and model misspecification appears in Subsection 6.2.2.

We are ready to begin presenting our work in detail. We highlight key background material and our paper's contributions in relation to it in Section 6.2. Our definition of mini-flash crashes is given and discussed in Section 6.3. Our agents and their beliefs are described in Section 6.4. We characterize the correct dynamics of the risky asset's price in Section 6.5. General results on what unfolds when our agents act as prescribed by Section 6.4 but prices actually move as in Section 6.5 are given in Section 6.6. Using the material in Section 6.6, a broad particular case of our model is investigated theoretically and numerically in Section 6.7. Our longer proofs are contained in Section 6.8 - 6.10.

6.2 Background & Contributions

In Section 6.2, we clarify our contributions and explain how they fit into the current literature. Subsections 6.2.1 and 6.2.2 contain the relevant discussions for our work on mini-flash crashes and model risk in optimal execution, respectively.

6.2.1 Mini-Flash Crashes

We already mentioned that existing theories on the causes of mini-flash crashes could be viewed as falling into one of five categories (see Section 6.1). Here are further details.

i) Human errors (and, relatedly, improper risk management) are among the most commonly cited causes of mini-flash crashes ([139], [167], [17]). The SEC claims that the majority of mini-flash crashes originate from such sources, in fact ([167]). When we read about *fat finger trades*, *rogue algorithms*, or *glitches* in the media, typically human errors are indirectly responsible. For example, due to a bug in the systems at the Tokyo Stock Exchange and a typo in a trade submitted by Mizuho Securities, the share price of the recruitment agency J-Com fell in minutes from ¥672,000 to ¥572,000 on December 8, 2005 ([2]).

ii) Mini-flash crashes may be caused by the rapid, endogenous formation of positive feedback loops ([8], [142], [106], [139], [131], [140]). As Johnson et al. put it,

> "Crowds of agents frequently converge on the same strategy and hence simultaneously flood the market with the same type of order, thereby generating the frequent extreme price-change events" ([142]).

A separate empirical study on the Flash Crash of May 6, 2010, specifically, determined that at its peak, "95% of the trading was due to endogenous triggering effects" ([106]).

iii) The nature of liquidity provision in modern markets is thought to cause some mini-flash crashes ([149], [96], [92], [116], [143], [115], [97]). Today, the majority of liquidity is provided by participants that are free from formal market-making obligations ([92]). In particular, they can instantly vanish, effectively taking one or both sides of the order book at some venue with them. A mini-flash crash can arise either directly as bid-ask spreads blow out or indirectly when a market order (of any size) tears through a nearly empty collection of limit orders. Such a phenomenon has been called *fleeting liquidity* and may have contributed to the occurrence of 38% of mini-flash crashes from 2006 - 2011 ([116]).

This proposed explanation is deeply intertwined with a crucial empirical observation: Mini-flash crashes occur in both high and low trading volume regimes.[1] For instance, the trading volume during the 30s mini-flash crash of "WisdomTree LargeCap" Growth Fund on November 27, 2012 was nearly eight times the average daily trading volume for this security ([17]). The empirical study by Florescu et al. offers extensive evidence that mini-flash crashes often occur during low trading volume periods as well.

Why modern liquidity providers might wish to briefly disappear at times is a separate issue. Broadly, the idea is that liquidity providers choose to pull back when they fear they will be adversely selected. Some suggest that the clearly erroneous trade regulations might discourage the submission of market-stabilizing

[1]Since market fragmentation may also contribute to local liquidity shortages (see (v)), it may be connected with this observation as well.

orders in the midst of a mini-flash crash ([92]). Gayduk and Nadtochiy propose a mechanistic theory: As trading frequencies increase, the very design of auction-style exchanges might ensure that markets become fragile and participants stop offering liquidity ([115]). Adverse selection fears are also stoked by genuine order flow toxicity, delays in consolidated quote feeds like the Security Information Processor, or activities by spoofers and other market manipulators ([5], [96], [97], [4]).

iv) Shocks to perceived fundamental values may lead to mini-flash crashes in some, albeit not most, cases ([143], [239]). We reiterate that these shocks must be perceived only: They may have no factual basis. For instance, a tweet sent on April 23, 2013 after a successful hack on the AP's Twitter account falsely claimed that President Obama was injured in a series of explosions at the White House. Within two minutes, $136 billion was erased from the S&P 500 Index ([147]).

v) Market fragmentation itself, as well as the current regulations concerning this issue, may give rise to some mini-flash crashes ([92], [116], [78]). In present-day markets, a particular security might be traded at a number of venues, and liquidity need not be uniformly distributed. This injects sophisticated considerations into the problem of optimal execution: How does one route an order to achieve the best possible price? The SEC introduced Rule 611, as well as various exceptions including intermarket sweep orders (ISOs), in an attempt to ensure that traders would receive the most favorable prices available across all venues ([16]). Some argue that, inadvertently, this regulation may have made matters worse. For instance, Dick posits a scenario in which a trader receives an inferior execution because Rule 611 only protects quotes at the top of the book ([92]). In their empirical analysis, Golub et al. find that most mini-flash crashes are initiated by aggressive ISO-submission ([116]).

Aspects of (i), (ii), and (iii) are reflected in our work. For example, the human error theory arises in each of the following ways:

a) Every agent believes that a mini-flash crash is a null event (see Remark VI.10). On the contrary, there are cases in which one will occur almost surely (see Lemmas VI.39 and VI.42).

b) Every agent thinks that his trades affect prices through specific temporary and permanent price impact coefficients (see Subsection 6.4.2). His estimates for these parameters might be wrong (see Section 6.5).

c) Every agent's trades may also indirectly impact prices by inducing others to make different decisions than they would otherwise (see Subsection 6.4.5 and Section 6.5). This potential effect is not modeled by our agents (see Subsection 6.4.2). More generally, even if we have a single agent in our setup trading with other unspecified market participants, the parameters in his fundamental value model might be inaccurate (see Subsection 6.4.2 and Section 6.5).

d) No agent revises the general class of his beliefs, admissible strategies, or objectives on our time horizon (see Subsections 6.4.2 - Subsection 6.4.4). In some cases, a mini-flash crash will not occur if this period is fairly short but will if it is too long (see Lemmas VI.28 and VI.34).

e) Every agent is averse to his position's apparent volatility risks (see Subsection 6.4.4). In some cases, there will be no mini-flash crash when our agents are sufficiently averse to these risks; otherwise, there will be one (see Lemmas VI.28 and VI.34).

f) Every agent has the opportunity to update the drift parameter in his price model based upon his observations (see Subsection 6.4.2). In some cases, a mini-flash crash will unfold because our agents are too easily persuaded to revise their priors (see Lemmas VI.28 and VI.34).

g) Every agent has a model for how prices are affected by the temporary impact of trades (see Subsection 6.4.2). In some cases, there will be a mini-flash crash if our agents sufficiently underestimate the role of aggregate temporary impact. No such disturbance will occur otherwise. Our agents may be more prone to induce mini-flash crashes in this way when there are many of them (see Lemmas VI.28 and VI.34).

Notice that some of our agents' human errors directly cause mini-flash crashes, though not all (see Lemma VI.28). We highlight this observation in Figures 6.1 - 6.3. Implicitly, the occasional absence of mini-flash crashes also agrees with (i). Despite the regularity of these disruptions on a market-wide basis, individual securities may rarely experience such an event. Similarly, traders' models and strategies do roughly achieve their intended goals much of the time, which we observe as well (see Lemma VI.28).

Several key ideas from the endogenous feedback loop theory are present in our paper. For example, if a mini-flash crash does occur, it almost surely does so because of "endogenous triggering effects." Specifically, our mini-flash crashes arise when some of our agents buy or sell at faster and faster rates, which they only do because they started trading more rapidly in the first place (see Section 6.5 and Lemma

VI.23). As predicted by this theory, some of our agents also "converge on the same strategy" during mini-flash crashes: In certain cases, the agents driving these events all buy or sell together with the same (exploding) growth rate (see Lemmas VI.39 and VI.42). Figures 6.5, 6.8, and 6.11 graphically illustrate this partial synchronization.

We do not explicitly model liquidity providers in our framework, as we view our agents as submitting market orders to a single venue (see Section 6.5). We still view our paper as reflecting (iii), at least in some sense, since our mini-flash crashes can be accompanied by both high and low trading volumes (see Corollary VI.15 and Lemmas VI.39 and VI.42). Visualizations of this point are provided in Figures 6.4, 6.6, 6.7, 6.9, 6.10, and 6.12.

The fundamental value shock theory is beyond the scope of our work. To study it, we could extend our model, say, by including a jump term in our specification of the actual price dynamics (see Section 6.5). Provided these jumps were almost surely finite, we suspect that they would not induce a mini-flash crash in the sense of our definition (see Section 6.3). This point is left for future work.

For the sake of tractability, we chose to model our agents as trading at a single venue (see Section 6.5). This puts the market fragmentation theory also beyond the scope of our paper. Especially since routing decisions are inextricably linked with optimal execution problems in practice, we hope to return to this topic in the future ([1]).

6.2.2 Model Risk & Optimal Execution

Problems in which agents make their decisions based upon misspecified models are well-studied in the economics and behavioral finance communities ([55], [60], [111], [35], [101], [24], [210], [100]). To the best of our knowledge, such an approach has not been directly pursued in the financial mathematics literature on optimal execution.

Much of the previous work in this area assumes that agents have complete and correct knowledge of all model parameters ([54], [19], [20], [22], [178], [113], [187], [45]). Others consider the possibility that their agents' models have the correct form; however, the agents must gradually learn the values of certain unobserved features ([21], [44], [75], [98], [110], [88]).

It is understood that anyone using the resulting strategies would be highly exposed to model (misspecification) risks. The concern is partially mitigated since methods including position limits, sensitivity analysis, Bayesian model averaging, the worst-case framework, and interpolations between the worst-case and classical setups may help to manage these issues. Agents that explicitly take model risks into account, say, by using one of these techniques, are typically called *ambiguity averse* ([73], [74], [76]).[2]

The idea with a control like a position limit is that although a model or strategy may never be perfect, their errors cannot cause ruinous damage. In practice, there are many related risk limits. Key differences among these variants tend to lie in what, specifically, is being limited in size and how its size limit is implemented ([171]). For example, the sizes of single positions, sector positions, market bets, market capitalization bets, and leverage might all be limited. The limits themselves might be inflexible constraints or appear as penalty functions.

Sensitivity analyses attempt to precisely measure how aspects of a strategy or its performance would change, if model assumptions are varied. Here, a model's parameters or probabilistic structure are often modified ([108], [107], [114]). If a strategy and its performance are found to be sufficiently stable, one might be somewhat assured that model risks are contained.

[2]Such agents appear throughout the literature, not just in the financial mathematics strand on optimal execution. The seminal book by Hansen & Sargent offers a comprehensive discussion ([120]). There have been many more recent developments as well (see [95], [203], [58], [51], [42], [40], [41], [43], [18], and the references therein).

Agents using the final three techniques above can be viewed as having several candidate models (not one). Alternatively, they can occasionally be interpreted as trying to reduce their exposure to Knightian (or unquantifiable) uncertainty ([150]). Possibly after assigning seemingly appropriate weights, such agents simultaneously measure the performance of their admissible strategies under all candidate models ([212], [89], [73]).

Despite the protection afforded by these methods, they do not offer complete inoculation against model risk.

Due to practical flaws in design or implementation, position limits may not always avert disaster. For example, the SEC found that in some cases, Merrill Lynch's controls allowed single orders to be placed with sizes that were over fifty times larger than a security's average daily trading volume ([17]). The SEC further claimed that these allegedly ineffective limits contributed to the onset of several mini-flash crashes, and Merrill Lynch was fined $12.5 million.

Certain types of sensitivity analysis, e.g., differentiating a strategy or its performance with respect to some parameter, may have shortcomings. For instance, they may be most useful when an agent's model is a slight perturbation of the actual price dynamics. Efficacy might be further lowered, if optimal strategies in one regime are only compared against optimal strategies in another (rather than studying how a single proposed strategy would perform under a new framework).

Safeguards provided by Bayesian model averaging, the worst-case framework, and interpolations between the worst-case and classical setups may be weakened if the agent's

i) candidate models all poorly represent the actual price dynamics,

ii) designated model weights are assigned inappropriately,

iii) or conceptions about future outcomes and their payoffs are mistaken.

Now, in much of the optimal execution literature with ambiguity averse agents, the sources of uncertainty are driven by Brownian motions, Poisson processes, or Poisson random measures, and the agent's candidate models are characterized by a suitable class of equivalent measures ([73], [74], [76]). Even when candidate models are allowed to be mutually singular, e.g., in the separate body of work on option pricing under volatility uncertainty ([165], [28]), they are *philosophically* similar, say, in the sense that they might have the same general form but differ in their parameter specifications. These points heighten the possibility of (i).

Weights corresponding to candidate models are typically determined in a Bayesian fashion or according to a chosen notion of distance from the candidate to the agent's reference model. (ii) may then arise, if either the agent's priors or belief metrics are not reflective of the actual price dynamics.

Historical examples of (iii) are abundant. Though not in the context of optimal execution, Taleb recounts an especially striking anecdote involving a casino ([226]). This organization put on a show which included a tiger. The firm insured against a variety of incidents but did not envision that the creature would attack its star performer. When this tragically occured, the casino lost $100 million and suffered one of its largest losses ever.

While model risks in optimal execution can never be entirely eliminated, these observations suggest that a new paradigm for managing them could be helpful. We hope that our general procedure, possibly applied together with existing techniques, might be such a paradigm. To clarify what we are introducing in the context of optimal execution, note the following alternative interpretation of our setup:

We have a single agent trading in a risky asset over a finite time horizon.

> In part, he has the beliefs and objectives described in Subsections 6.4.2 - 6.4.4; however, he is also concerned about model misspecification risks. Before he begins trading, he wishes to have a more robust understanding of how the strategy he derives in Subsection 6.4.5 might perform. To get this, he imagines a new plausible way that the price might evolve and tests his strategy's performance in this scenario. He first hypothesizes that there might be other market participants who stumbled upon the basics of his strategy and might be planning to use these ideas. He is then led to consider the possibility that the actual price dynamics are as given in Section 6.5. He studies what might unfold in Sections 6.6 - 6.7. He concludes that since it seems like his original strategy might lead to mini-flash crashes and devastating losses at times, he would like to reconsider his trading plans.

In short, instead of emphasizing *mathematical* similarity when checking his strategy's performance in additional models, we could view our agent as emphasizing his *human* similarity with other market participants. By doing so, he seems to test his ideas in alternative settings which are both plausible and yield a different perspective on his model risks (compared to the insights offered by more traditional approaches). Observe that the idea that our agents, real or fictitious, individually solve their optimal execution problems and are not in a classical equilibrium state is crucial here.

The concept that strategy replication among market participants may significantly affect the future and, hence, may bring unforeseen risks is not limited to the context of optimal execution. For example, there is growing concern that the dramatic rise in index investing may have unanticipated, detrimental effects on the broader economy

([61], [99], [29], [30]).

Also, when we consider our setup from this new viewpoint, the consistency issues which may arise in conjunction with some of the human errors in our framework are not worrisome (see (a) - (g) in Subsection 6.2.1). After all, our agent is deliberately falsifying his beliefs to better discern his model risk exposure.

Even if we retain our initial finite population system view, we feel that these concerns may not be significant. While we could directly attempt to fit our framework into one of the modern equilibrium notions in the model misspecification literature, there may not be a need to do so: It appears that there are simple, practically-oriented reasons why the apparent issues might naturally come about in our setting.

First, we think of our agents as having the opportunity to witness just a single realization of the price, meaning that each agent believes that null events (including the price path itself) must occur. This seems to reflect the non-stationarity of markets: Models, parameters, and strategies which work quite well in one period may fail in the next.

Second, in practice, agents do not necessarily notice all of the ways in which their models are wrong. If they do, they may not want or be able to fix them. These behavioral arguments seem especially valid over the short time horizons that we consider and are supported by general observations from both the philosophy of science and psychology communities ([79], [23], [118], [222], [169]). For a specific example verifying these claims, recall the circumstances which engulfed Knight Capital on August 1, 2012: A bug arose in a critical piece of software. It has been alleged that the firm did not detect the glitch themselves; rather, they only became aware of it after being notified by the New York Stock Exchange. Supposedly, it then took 30 to 45 minutes for the firm to implement corrections, leading to a \$440 million loss and Knight Capital's subsequent acquisition by Getco LLC ([198]).

6.3 Mini-Flash Crashes

In Section 6.3, we introduce and discuss our definition of mini-flash crashes.

Definition VI.1. We say that a *mini-flash crash* occurs, if the risky asset's price tends to either $+\infty$ or $-\infty$ on our time horizon.

For now, while this definition communicates our broad notion, its details are fairly vague.

We say more about our time horizon and price in Sections 6.4 - 6.5. The former is finite and deterministic, while the latter is a particular stochastic process.

We precisely describe the sense in which the price explodes and at what kinds of times[3] mini-flash crashes can occur in Sections 6.6 - 6.7. Roughly, we analyze the occurrence of mini-flash crashes pathwise. In the cases that we consider, they happen (or not) almost surely; however, their direction is random. If a mini-flash crash unfolds, it does so at a deterministic time; yet, none of our agents have enough information to compute this time or even know that a mini-flash crash is imminent.

We said nothing about the classification of unbounded price oscillations as well. In the scenarios that we investigate, unbounded price oscillations occur with probability zero (see Sections 6.6 - 6.7), but this may be an artifact of our technical choices. Such an oscillation appears to reflect the practical duration of mini-flash crashes, and we hope to explore this point in a future work.

Clearly, there are many reasons why mini-flash crashes would be non-existent, if Definition VI.1 were used in practice. For instance, the SEC has instituted the "Limit Up-Limit Down Mechanism" to temporarily suspend trading in individual securities

[3] We thank Shige Peng for this observation.

whose prices escape certain upper and lower bounds in specified short periods ([15]). Market-wide circuit breakers might be employed as well, which temporarily halt all trading when the S&P 500 Index declines sufficiently in a single trading day ([15]).

Our intuitive justification for this approximation is four-fold: First, though finite, price swings during a mini-flash crash can be quite extreme and shocking (see Section 6.1 and Subsection 6.2.1). Second, in our setting, we roughly view that trading would be suspended just before the occurrence of the event in Definition VI.1: A mini-flash crash is more appropriately understood to be the local behavior of the price near such a disruption. Third, since there are cases in which prices explode almost surely in our framework, Definition VI.1 avoids seemingly more arbitrary cut-offs that might have been necessary, if we included finite disturbances (see Sections 6.1 and 6.7). Finally, when we view our setup from the model risk-averse single agent perspective explained in Subsection 6.2.2, it might be reasonable to hypothesize that our agent would consider the possibility of exploding prices, even as only a limiting case which must be averted.

6.4 Agents

In Section 6.4, we describe our agents and their individual optimal execution problems. Important preliminary details are given in Subsection 6.4.1. We present our agents' models and beliefs in Subsection 6.4.2. Each agent's admissible strategies are characterized in Subsection 6.4.3. We discuss the agents' objectives in Subsection 6.4.4. We prove Lemma VI.9, the main result of Section 6.4, in Subsection 6.4.5. Lemma VI.9 prescribes optimal strategies for our agents given their beliefs and preferences. Our agents attempt to trade according to these plans on our time horizon. When he does so, each agent believes that a mini-flash crash will occur with probability zero.

6.4.1 Preliminaries

We consider a population of N agents: Agent 1,..., Agent N. The intuition underlying our agent's models suggests that N should be interpreted as a large number (see Subsection 6.4.2). Mathematically, its qualitative size does not matter (see Lemma VI.23).

There is a special nonnegative parameter ν_j^2 associated to Agent j (see Subsection 6.4.2).

Definition VI.2. If $\nu_j^2 > 0$, then we call Agent j an *uncertain* agent. If $\nu_j^2 = 0$, then we call Agent j a *certain* agent.

The reason for choosing these particular words will become clear in Subsection 6.4.2, and the distinction between these two types of agents will be crucial throughout the rest of the paper. For now, we assume that there are $K \in \{0, \dots, N\}$ uncertain agents, namely, Agents 1 through K. A critical role is played by the value of K (see Lemma VI.23).

All agents attempt to trade in a single risky asset over a time horizon $[0, T]$. Our arguments proceed as long as T is deterministic and finite; however, our rationale is reasonable only when this period is short, say, no more than 1 day (see Subsection 6.4.2).

6.4.2 Models

Our agents trade continuously by optimally selecting a trading rate from a particular class of admissible strategies. To motivate our specifications of their choices and objectives, we first define their models and beliefs.

All trades submitted at time t are executed immediately at the price S_t^{exc}. At each time t, every agent observes the correct value of S_t^{exc}. No agent knows the true dynamics of the stochastic process S^{exc}, though.

Instead, prior to $t = 0$, Agent j has developed a model S_{j,θ_j}^{exc} for S^{exc}. S_{j,θ_j}^{exc} evolves on

$$\left(\Omega_j, \mathcal{F}_j, \{\mathcal{F}_{j,t}\}_{0\leq t\leq T}, P_j\right), \tag{6.1}$$

a filtered probability space satisfying the usual conditions.[4] Every agent models S^{exc} on a different probability space, despite the fact that their observations of S^{exc} will be identical. Our point is that Agent j interprets his observations as a sample path of his individual model for S^{exc}, which may be unrelated to the process that Agent k uses to interpret the same data.

The space (6.1) comes equipped with W_j, an $\mathcal{F}_{j,t}$-Wiener process under P_j. There is also an $\mathcal{F}_{j,0}$-measurable random variable β_j, which is independent of W_j and normally distributed with mean μ_j and variance ν_j^2 under P_j.

Recalling Definition VI.2, we see that Agent j is *certain* if he believes that he knows the correct value of β_j at $t = 0$. Otherwise, he is *uncertain*. Regardless of whether he is certain or uncertain in this sense, we will soon see that Agent j can always be viewed as certain about many things, e.g, he will not change the form of his models, objectives, or admissible strategies on $[0, T]$.[5]

Definition VI.3. Following ([21]), Agent j defines an $\mathcal{F}_{j,t}$-adapted process S_j^{unf} by

$$S_{j,t}^{unf} = S_{j,0} + \beta_j t + W_{j,t}, \qquad t \in [0, T]. \tag{6.2}$$

[4]From a technical perspective, we will soon see that there is no need to introduce the filtration $\{\mathcal{F}_{j,t}\}$. It would be equivalent to work with $\left\{\mathcal{F}_{j,t}^{unf}\right\}$ (see Subsections 6.4.3 and 6.4.5). The basic observation is that Agent j believes he can correctly reformulate his original optimal execution problem with partial information as one with complete information. Keeping the first problem seems to help motivate our setup.

[5]From this perspective, one might partially connect our work on mini-flash crashes to explanations of longer term financial bubbles based on overconfident investors ([211]).

$S_{j,t}^{unf}$ is Agent j's estimate of the *unaffected* or *fundamental* price of the risky asset at time t. The drift term represents the price pressure that Agent j believes will arise due to the trades of (other) institutional investors. Agent j approximates the average behavior of uninformed or noise traders using the Brownian term.[6]

After a fashion, Agent j believes that he can compute $S_{j,t}^{unf}$ at t (see Subsection 6.4.3). Implicitly, he believes that his observations of $S_{j,t}^{unf}$ will be independent of his trading decisions. Agent j knows which deterministic constant $S_{j,0}$ he has selected in (6.2). Unless $\nu_j^2 = 0$, he does not assume that he can determine the realized values of β_j or $W_{j,t}$. Instead, Agent j will attempt to learn the value of β_j by computing its expectation conditional on his accumulated observations as time passes (see Subsection 6.4.5).

Intuitively, Agent j's selection of (6.2) makes the most sense when N is large and T is short. Notice that Agent j makes no attempt to precisely estimate the number of other market participants, nor their individual goals or beliefs. That he believes he cannot improve the predictive accuracy of (6.2) by doing so appears to suggest that the population of traders is of sufficient size.[7] In practice, many securities' prices must be positive.[8] Together with the fact that real drifts and volatilities are non-constant, (6.2) only seems even potentially plausible over short periods.

We now are ready to discuss S_{j,θ_j}^{exc}.

Definition VI.4. Let $\theta_{j,t}$ denote Agent j's trading rate at time t (see Subsection

[6]Almgren & Lorenz provide further details regarding the interpretation and limitations of (6.2) ([21]). A possible extension of our work could replace (6.2) with one of the more recent models considered in the literature on optimal trading problems with a learning aspect ([21], [98], [75], [113], [187], [110], [88]).

[7]Alternatively, one could argue that there are only a few agents, all of whom are effectively hidden from one another; however, the securities for which our framework seems most reasonable would probably be traded by a large population anyway.

[8]Certain commodities have traded at negative prices ([11]).

6.4.3). He defines S^{exc}_{j,θ_j} as the $\mathcal{F}_{j,t}$-adapted process

$$S^{exc}_{j,\theta_j,t} = S^{unf}_{j,t} + \eta_{j,per}\int_0^t \theta_{j,s}\,ds + \frac{1}{2}\eta_{j,tem}\theta_{j,t}, \qquad t \in [0,T]\,. \tag{6.3}$$

Agent j has chosen the deterministic positive constants $\eta_{j,per}$ and $\eta_{j,tem}$ in (6.3) prior to time $t = 0$.

There are two primary rationales behind (6.3). First, Agent j could be viewed as taking into account his own effects on the execution price via an Almgren-Chriss reduced-form model ([20], [19], [22]). $\eta_{j,per}$ would denote Agent j's estimate for his permanent price impact parameter, while he would approximate his temporary price impact parameter with $\eta_{j,tem}$. There is an alternative explanation in which Agent j believes he submits market orders to a limit order book with certain characteristics. We present further details for both viewpoints in Section 6.5.

While Agent j can use his prior for β_j, as well as his observations, to improve his estimate for the realized value of β_j, all model parameters in (6.2) and (6.3) are fixed (see Subsection 6.4.3). He cannot change the form of these models either, e.g., by making β_j time-dependent in (6.2) or including a transient impact term in (6.3). The idea is that T is short and, in practice, the time scale for developing an appropriate class of models for trading some instrument is often much longer than the time scale for revising parameters to better reflect current market conditions.

6.4.3 Admissible Strategies

Agent j selects his trading rate θ_j from $\mathcal{A}_j$, a class of admissible strategies that we will now define precisely.

Recall that Agent j does not observe the realizations of either β_j or W_j. Hence, it would not make sense for Agent j's trading rate to be $\mathcal{F}_{j,t}$-adapted. Agent j does watch S^{exc}, though, which he interprets as S^{exc}_{j,θ_j}. Working with the filtration generated by S^{exc}_{j,θ_j} is somewhat cumbersome, as Agent j believes that it depends on his choice of trading rate. The key is to notice that when Agent j selects a continuous trading rate adapted to $\left\{\mathcal{F}^{unf}_{j,t}\right\}$, the filtration generated by S^{unf}_j, he believes that this filtration describes the same flow of information as his execution price observations. This is advantageous, as Agent j thinks that $\left\{\mathcal{F}^{unf}_{j,t}\right\}$ is independent of his trading decisions (see Subsection 6.4.2).

Intuitively, the idea is that at each time t, Agent j observes the correct value of S^{exc}_t. He views this value as the realization of $S^{exc}_{j,\theta_j,t}$. Using his knowledge of his past trading rate, he determines $S^{unf}_{j,t}$ as in (6.3). Agent j thinks that these steps can be effectively taken all at once due to the (perceived) continuity of each process involved in the calculations.

Agent j also believes his trades suffer from transaction costs due to both temporary and permanent price impact (see (6.3)). It seems reasonable to assume that he would never adopt a strategy that he thought would saddle him with infinite costs. As temporary impact induces a quadratic cost, we specify that he can only choose a strategy that satisfies

$$E^{P_j}\left[\int_0^T \theta_{j,t}^2\, dt\right] < \infty. \tag{6.4}$$

Costs arising from permanent impact do not depend on Agent j's trading rate, if his terminal inventory is deterministic. In fact, for reasons discussed in Subsection 6.4.4, we specify that Agent j's terminal inventory must be zero, i.e., Agent j solves an optimal liquidation problem.[9]

Formalizing these comments leads to the following definition.

Definition VI.5. Let $\mathcal{A}_j$ be the space of $\mathcal{F}^{unf}_{j,t}$-adapted processes θ_j such that $\theta_{j,\cdot}(\omega)$

[9] Why would Agent j produce the estimate $\eta_{j,per}$? After all, he believes that its value will not affect his trading decisions. Roughly, we feel that Agent j might have such an approximation for business purposes, e.g., he may hope to accurately forecast P&L, even if he believes some components are uncontrollable. Unbeknownst to Agent j, $\eta_{j,per}$ is quite important for additional reasons (see Subsection 6.7.1).

is continuous on $[0, T]$ for P_j-almost every $\omega \in \Omega_j$, (6.4) holds, and

$$x_j + \int_0^T \theta_{j,t}\, dt = 0 \quad P_j - \text{a.s.} \tag{6.5}$$

For any $\theta_j \in \mathcal{A}_j$, we define the process $X_j^{\theta_j}$ by

$$X_{j,t}^{\theta_j} = x_j + \int_0^t \theta_{j,s}\, ds. \tag{6.6}$$

$X_j^{\theta_j}$ is our notation for Agent j's inventory: When he trades according to θ_j, Agent j holds $X_{j,t}^{\theta_j}$ shares of the risky asset at time t. In particular, we could also write (6.5) as

$$X_{j,T}^{\theta_j} = 0 \quad P_j - \text{a.s..}$$

In agreement with our intuition, the process $X_j^{\theta_j}$ is $\mathcal{F}_{j,t}^{unf}$-adapted due to our construction of $\mathcal{A}_j$.

6.4.4 Objective Functions

Agent j would like to trade such that, on average, his realized trading revenue will be as high as possible. He is also concerned about the various risks he might encounter while trading and hopes to take some of these into account. Since Agent j proxies S^{exc} with S_{j,θ_j}^{exc}, he believes that the expected revenue corresponding to $\theta_j \in \mathcal{A}_j$ is given by

$$E^{P_j}\left[- \int_0^T \theta_{j,t}\, S_{j,\theta_j,t}^{exc}\, dt \right]. \tag{6.7}$$

Now Agent j must consider how to manage several risks. First, there are volatility risks associated with delayed liquidation. Since he uses (6.2) and (6.3), he believes that these can be quantified by

$$E^{P_j}\left[-\frac{\kappa_j}{2} \int_0^T \left(X_{j,t}^{\theta_j} \right)^2 dt \right]. \tag{6.8}$$

In (6.8), Agent j selects the deterministic risk aversion parameter $\kappa_j > 0$ based upon his appetite. On the other hand, Agent j presumably believes that as he observes the execution price's path, he can better estimate β_j's realized value. He might then think that he is more likely to regret earlier trades than later trades. A simple, though admittedly ad-hoc, way that Agent j could adjust for this risk is to artificially lower κ_j. Similarly, it might be possible for him to partially account for his other risks including those arising from model misspecification with such an approach. In fact, in a slightly different setting, Jaimungal et al. show the equivalence between certain forms of ambiguity aversion and quadratic inventory penalties ([73]).

This discussion suggests the following objective for Agent j.

Definition VI.6. Agent j's objective is to maximize

$$E^{P_j}\left[-\int_0^T \theta_{j,t}\, S^{exc}_{j,\theta_j,t}\, dt - \frac{\kappa_j}{2}\int_0^T \left(X^{\theta_j}_{j,t}\right)^2 dt\right] \tag{6.9}$$

over $\theta_j \in \mathcal{A}_j$.

As mentioned in Subsection 6.4.3, (6.9) motivates our requirement that Agent j must liquidate by T (see (6.5)). Although the permanent impact term in (6.9) disappears, regardless of the deterministic value of $X^{\theta_j}_{j,T}$, non-zero values of $X^{\theta_j}_{j,T}$ could perversely incentivize Agent j via (6.8). For example, if Agent j started with a large inventory and needed a larger terminal inventory, he might be inclined to pay unnecessary round-trip (sell early/buy later) costs induced by temporary impact.

In Lemma VI.9's proof, we see that (6.9) can be equivalently formulated as the following complete information problem: Maximize

$$E^{P_j}\left[\int_0^T X^{\theta_j}_{j,t}\, E^{P_j}\left[\beta_j \middle| \mathcal{F}^{unf}_{j,t}\right] dt - \frac{\eta_{j,tem}}{2}\int_0^T \theta^2_{j,t}\, dt - \frac{\kappa_j}{2}\int_0^T \left(X^{\theta_j}_{j,t}\right)^2 dt\right]$$

over $\theta_j \in \mathcal{A}_j$. It can also be thought of as an optimal tracking problem, in which Agent j must minimize

$$E^{P_j}\left[\frac{1}{2}\int_0^T \left(X^{\theta_j}_{j,t} - \frac{E^{P_j}\left[\beta_j \middle| \mathcal{F}^{unf}_{j,t}\right]}{\kappa_j}\right)^2 dt + \frac{\eta_{j,tem}}{2\kappa_j}\int_0^T \theta^2_{j,t}\, dt\right].$$

Variants of both the former and the latter have been previously investigated, although not, to the best of our knowledge, with our intentions ([113], [34]).

6.4.5 Results

Before proving Lemma VI.9, we introduce the following notation. It will be useful throughout the paper.

Definition VI.7. We define $\tau_j(\cdot)$ by

$$\tau_j(t) \triangleq \sqrt{\frac{\kappa_j}{\eta_{j,tem}}}(T-t), \quad t \in [0,T].$$

Remark VI.8. For our purposes, the key point in Definition VI.7 is that τ_j strictly decreases to 0 as $t \uparrow T$.

Lemma VI.9. *(6.9) has a unique*[10] *solution* $\theta_j^\star \in \mathcal{A}_j$. *When* $\omega \in \Omega_j$ *is chosen such that* $W_{j,\cdot}(\omega)$ *is continuous on* $[0,T]$, $X_j^{\theta_j^\star}(\omega)$ *satisfies the linear ODE*

$$\begin{aligned}
\theta_{j,t}^\star(\omega) &= -\sqrt{\frac{\kappa_j}{\eta_{j,tem}}}\coth(\tau_j(t))\, X_{j,t}^{\theta_j^\star}(\omega) \\
&\quad + \frac{\tanh(\tau_j(t)/2)\left[\mu_j + \nu_j^2\left(S_{j,t}^{unf}(\omega) - S_{j,0}\right)\right]}{\sqrt{\eta_{j,tem}\kappa_j}\left(1+\nu_j^2 t\right)}, \quad t \in (0,T) \\
X_{j,0}^{\theta_j^\star}(\omega) &= x_j.
\end{aligned} \tag{6.10}$$

Remark VI.10. In conjunction with Subsections 6.4.2 - 6.4.3, Lemma VI.9 implies that Agent j believes that a mini-flash crash is a null event. More precisely, under his setup, S_{j,θ_j}^{exc}, hence S^{exc}, will remain finite on $[0,T]$ P_j-a.s. Of course, he believes that this will be true of $X_j^{\theta_j^\star}$ and $\theta_j^\star$ as well.

Remark VI.11. Agent j believes that (6.10) characterizes his optimal trading rate almost surely. Therefore, it seems reasonable to view that he would *always* attempt to implement this strategy: He thinks it is nearly impossible for this approach to be flawed, after all.

[10]Here, uniqueness holds up to $dP_j \otimes dt$-a.s. equality on $\Omega_j \times [0,T]$.

Remark VI.12. The first term in (6.10) arises from our constraint that Agent j must liquidate by the terminal time (see (6.5)). In fact, the weighting factor

$$-\sqrt{\frac{\kappa_j}{\eta_{j,tem}}}\coth\left(\tau_j\left(t\right)\right)$$

tends to $-\infty$ as $t \uparrow T$. Intuitively, the reason that Agent j believes that $X_{j,t}^{\theta_j^\star}$ and $\theta_{j,t}^\star$ remain finite as $t \uparrow T$ is that $X_{j,t}^{\theta_j^\star}$ tends very rapidly to zero.

Agent j thinks that he learns about β_j's realized value over time, which is captured by the second term in (6.10) since

$$E^{P_j}\left[\beta_j \middle| \mathcal{F}_{j,t}^{unf}\right] = \frac{\mu_j + \nu_j^2\left(S_{j,t}^{unf} - S_{j,0}\right)}{1 + \nu_j^2 t} \quad P_j - \text{a.s.} \tag{6.11}$$

([162]). The factor

$$\frac{\tanh\left(\tau_j\left(t\right)/2\right)}{\sqrt{\eta_{j,tem}\kappa_j}}$$

is bounded by $1/\sqrt{\eta_{j,tem}\kappa_j}$ and tends to zero as $t \uparrow T$.

The second term may either dampen or amplify the effects of the first. Agent j believes that the weighting factors reflect that his need to liquidate must eventually overwhelm his desire to profit by trading in the direction of the risky asset's drift.

Remark VI.13. As anticipated, Agent j's permanent impact parameter estimate $\eta_{j,per}$ is absent in (6.10) (see Subsection 6.4.3).

Remark VI.14. Lemma VI.9's proof has five steps.

First, we introduce an auxiliary problem in which Agent j can select a trading rate from a larger class of admissible strategies. Our original formulation did not consider these, as they may not be aligned with the intuition underlying our framework. For instance, some of them suggest that Agent j could peak into the future or that he might *knowingly* select a trading rate that would cause S_{j,θ_j}^{exc} to explode.

We then show that Agent j does not believe that he can benefit from the auxiliary problem's new informational structure. This part of the argument uses (6.11), as well as the Vitali convergence theorem and uniform integrability.

The third step is to find a suitable complete information equivalent for Agent j's auxiliary problem. We have effectively discussed this in Subsection 6.4.4. The idea is to use integration by parts and an innovation process.

Next, we use a result from Bank et al. to determine the unique solution to our auxiliary problem ([34]). The introduction of our auxiliary problem was motivated by this step. Admittedly, this means we use a tool which seems to be far more powerful than our problem demands. For instance, the result of Bank et al. applies to a broad class of optimal tracking problems with a non-Markovian target, while (6.11) suggests that we track a Markovian one (see Subsection 6.4.4). We still adopt this approach, as it may allow us to extend our work in the future.

We conclude by demonstrating that the trading rate identified in the previous step is actually in our original set of admissible strategies. Much of the work to prove that the trading rate is adapted to $\left\{\mathcal{F}_{j,t}^{unf}\right\}$ comes from our second step, while the remaining ideas are taken care of by Bank et al. ([34]).

Proof. See Subsection 6.8.1. □

Corollary VI.15. *If $\nu_j^2 = 0$, then $X_j^{\theta_j^\star}$ does not depend on S_j^{unf}. In particular, it is deterministic and satisfies the linear ODE*

$$\theta_{j,t}^\star = -\sqrt{\frac{\kappa_j}{\eta_{j,tem}}}\coth\left(\tau_j\left(t\right)\right)X_{j,t}^{\theta_j^\star} + \frac{\mu_j \tanh\left(\tau_j\left(t\right)/2\right)}{\sqrt{\eta_{j,tem}\kappa_j}}, \quad t\in(0,T)$$
$$X_{j,0}^{\theta_j^\star} = x_j. \tag{6.12}$$

Remark VI.16. Corollary VI.15 confirms that there are significant differences between our certain and uncertain agents, as expected: If Agent j feels completely certain of

β_j's realized value, he would not glean profitable information and modify his trades based upon his observations of the execution price (see Subsections 6.4.3 - 6.4.4). Mathematically, it is also especially evident from (6.37) and (6.38).

Proof. This is immediate from Lemma VI.9. □

6.5 Execution Price

In Section 6.5, we specify how S^{exc} actually evolves. While each agent observes the same realized path of this process, in general, no agent knows the correct dynamics.[11] An agent's trading decisions are entirely determined by his beliefs, preferences, and observations of a single realized path of S^{exc} (see Lemma VI.9).

Let

$$\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{0 \leq t \leq T}, \tilde{P}\right)$$

be a filtered probability space satisfying the usual conditions. The space is equipped with an $\tilde{\mathcal{F}}_t$-Wiener process under $\tilde{P}$, which we denote by $\tilde{W}$. We also have the following deterministic real constants:

$$\tilde{\beta}, \quad S_0, \quad \tilde{\eta}_{1,per}, \ldots, \tilde{\eta}_{N,per}, \quad \text{and} \quad \tilde{\eta}_{1,tem}, \ldots, \tilde{\eta}_{N,tem}.$$

$\tilde{\beta}$ can be arbitrary; however, the remaining constants are strictly positive.

Definition VI.17. The true execution price S^{exc} under $\tilde{P}$ is the $\tilde{\mathcal{F}}_t$-adapted process

$$S_t^{exc} = S_0 + \tilde{\beta} t + \sum_{i=1}^{N} \tilde{\eta}_{i,per} \left(X_{i,t}^{\theta_i^\star} - x_i \right) + \frac{1}{2} \sum_{i=1}^{N} \tilde{\eta}_{i,tem} \theta_{i,t}^\star + \tilde{W}_t, \qquad t \in [0, T]. \quad (6.13)$$

(6.13) can be viewed as a multi-agent extension of the Almgren-Chriss model ([20], [19], [22]). Models of this form, particularly when the $\tilde{\eta}_{j,tem}$'s ($\tilde{\eta}_{j,per}$'s) are all identical, have been applied in the context of predatory trading ([69]).

[11] There is a single trivial case where this is not true. If $N = 1$, $\tilde{\beta} = \beta$, $\nu_1^2 = 0$, $\tilde{\eta}_{1,tem} = \eta_{1,tem}$, and $\tilde{\eta}_{1,per} = \eta_{1,per}$, our lone agent's model would be exactly right.

From this perspective, $\tilde{\eta}_{j,per}$ and $\tilde{\eta}_{j,tem}$ are the *correct* values of Agent j's permanent and temporary price impact parameters, respectively. We allow these quantities to have arbitrary relationships to Agent j's corresponding *estimates* $\eta_{j,per}$ and $\eta_{j,tem}$. For instance, Agent j might underestimate his permanent impact ($\eta_{j,per} < \tilde{\eta}_{j,per}$) but perfectly estimate his temporary impact ($\eta_{j,tem} = \tilde{\eta}_{j,tem}$). Similarly, Agent j's prior β_j for the *correct* drift $\tilde{\beta}$ may be accurate or severely mistaken. Comparing our descriptions of S^{exc}_{j,θ_j} in (6.3) and S^{exc} in (6.13), we see that Agent j proxies each term in (6.13) as follows:

$$\begin{aligned}
\eta_{j,per}\left(X^{\theta_j^\star}_{j,t} - x_j\right) &\quad\longleftrightarrow\quad \tilde{\eta}_{j,per}\left(X^{\theta_j^\star}_{j,t} - x_j\right) \\
\frac{1}{2}\eta_{j,tem}\theta^\star_{j,t} &\quad\longleftrightarrow\quad \frac{1}{2}\tilde{\eta}_{j,tem}\theta^\star_{j,t} \\
S_{j,0} + \beta_j t + W_{j,t} &\quad\longleftrightarrow\quad S_0 + \tilde{\beta}t + \sum_{i\neq j}\tilde{\eta}_{i,per}\left(X^{\theta_i^\star}_{i,t} - x_i\right) + \frac{1}{2}\sum_{i\neq j}\tilde{\eta}_{i,tem}\theta^\star_{i,t} + \tilde{W}_t.
\end{aligned}$$

Heuristically, we could also interpret (6.13) through the lens of a single order book. This connection was observed by Kallsen & Muhle-Karbe ([144]). The process

$$S_0 + \tilde{\beta}t + \tilde{W}_t$$

would be viewed as the fundamental price, while the sum of the fundamental price and the permanent impact terms

$$S_0 + \tilde{\beta}t + \tilde{W}_t + \sum_{i=1}^{N}\tilde{\eta}_{i,per}\left(X^{\theta_i^\star}_{i,t} - x_i\right)$$

would be the reference price. We would set the $\tilde{\eta}_{j,tem}$'s to a single value, and do the same for the $\tilde{\eta}_{j,per}$'s. Our agents would submit market orders, and only the net agent order flow would be executed in the book (remaining orders would be matched together). All agents would receive the same average execution price at each time. The bid-ask spread would be infinitesimally small, while the book would be infinitely resilient and block-shaped with height $1/\tilde{\eta}_{j,tem}$. That is, agents would trade in an Obizhaeva-Wang book which instantly recovers to the reference price after each execution (no transient impact) ([178]) .

6.6 General Results

When our agents implement the strategies that they believe are optimal (see Lemma VI.9) but S^{exc} has the dynamics in (6.13), what happens? The goal of Section 6.6 is to offer some general answers to this question.

To simplify our presentation, we begin by introducing and analyzing additional notation (see Definition VI.19 and Lemma VI.21). We find that our agents' inventories and trading rates evolve according to a particular ODE system with stochastic coefficients (see Lemma VI.23). Under certain conditions, the system can have a singular point (see Lemma VI.24). For convenience, we study what unfolds when this singular point is of the first kind (see Lemma VI.27). We also examine the case in which there is no singular point (Lemma VI.28). Due to tractability issues, in order to determine whether or not mini-flash crashes arise, we consider a particular, though broad, class of examples (see Remark VI.30). While we present these findings in Section 6.7, we provide a high-level summary of them in Theorem VI.31. In particular, we see that in some cases, mini-flash crashes occur $\tilde{P}$-a.s. at the first singular point of our system. Again, our agents still believe that a mini-flash crash is a null event.

First, observe that our assumptions in Section 6.5 do not affect our certain agents' trading decisions (see Corollary VI.15). It remains to characterize our uncertain agents' strategies.

We will have an even mix of deterministic and stochastic maps. In what follows, we always explicitly indicate ω-dependence to distinguish between the two. Our equations are solved pathwise, so we do not encounter probabilistic concerns.

Notation VI.18. Fix $\omega \in \tilde{\Omega}$ such that $\tilde{W}_{\cdot}(\omega)$ has a continuous path.

Definition VI.19. Define the maps

$$\begin{aligned} \Phi_i : [0,T] &\longrightarrow \mathbb{R} \\ A : [0,T] &\longrightarrow M_K(\mathbb{R}) \\ B : [0,T) &\longrightarrow M_K(\mathbb{R}) \\ C(\cdot,\omega) : [0,T] &\longrightarrow \mathbb{R}^K \end{aligned}$$

by

$$\begin{aligned} \Phi_i(t) &\triangleq \frac{\tanh(\tau_i(t)/2)\,\nu_i^2}{\sqrt{\eta_{i,tem}\kappa_i}\,(1+\nu_i^2 t)} \\ A_{ik}(t) &\triangleq \begin{cases} 1 - \frac{1}{2}(\tilde{\eta}_{i,tem} - \eta_{i,tem})\,\Phi_i(t) & \text{if } i = k \\ -\frac{1}{2}\tilde{\eta}_{k,tem}\Phi_i(t) & \text{if } i \neq k \end{cases} \\ B_{ik}(t) &\triangleq \begin{cases} (\tilde{\eta}_{i,per} - \eta_{i,per})\,\Phi_i(t) - \sqrt{\frac{\kappa_i}{\eta_{i,tem}}}\coth(\tau_i(t)) & \text{if } i = k \\ \tilde{\eta}_{k,per}\Phi_i(t) & \text{if } i \neq k \end{cases} \\ C_i(t,\omega) &\triangleq \Phi_i(t)\left[\frac{\mu_i}{\nu_i^2} + (S_0 - S_{i,0}) + \tilde{\beta}t - \sum_{\substack{k\leq K\\ k\neq i}} \tilde{\eta}_{k,per}x_k - x_i(\tilde{\eta}_{i,per} - \eta_{i,per})\right. \\ &\qquad \left. + \sum_{k>K} \tilde{\eta}_{k,per}\left(X_{k,t}^{\theta_k^\star} - x_k\right) + \frac{1}{2}\sum_{k>K}\tilde{\eta}_{k,tem}\theta_{k,t}^\star + \tilde{W}_t(\omega)\right]. \end{aligned}$$

Here, $i \in \{1,\ldots,K\}$.

Remark VI.20. Observe that we can now write the dynamics in (6.10) as

$$\theta_{j,t}^\star(\omega) = -\sqrt{\frac{\kappa_j}{\eta_{j,tem}}}\coth(\tau_j(t))\,X_{j,t}^{\theta_j^\star}(\omega) + \Phi_j(t)\left[\frac{\mu_j}{\nu_j^2} + \left(S_{j,t}^{unf}(\omega) - S_{j,0}\right)\right]$$

when Agent j is uncertain.

We frequently reference various easy properties of the functions in Definition VI.19. We collect these below for convenience.

Lemma VI.21. *Fix $j \in \{1, \ldots, K\}$. We have the following:*

i) Φ_j is a strictly decreasing nonnegative function on $[0, T]$ with $\Phi_j(T) = 0$.

ii) The entries of A are analytic on $[0, T]$ and $A(T) = I_K$.

iii) If $\det A$ has a root on $[0, T]$, we can find the smallest one which we denote by t_e. In this case, $t_e < T$ and the zero of $\det A$ at t_e is of finite multiplicity.

iv) The entries of B are analytic on $[0, T)$ but

$$\lim_{t \uparrow T} B_{jj}(t) = -\infty.$$

v) $C(\cdot, \omega)$'s entries are continuous on $[0, T]$.

Proof. Parts (i) and (ii) are clear. After recalling that the zeros of an analytic function are isolated and of finite multiplicity, we get (iii) from (ii). The singularity in B_{jj} at T arises from the coth term, yielding (iv). Corollary VI.15 and our choice of ω give (v). $\square$

Definition VI.22. When $\det A$ has a root on $[0, T]$, we let t_e denote the smallest one (see Lemma VI.21).

Up to some deterministic time, the uncertain agents' inventories evolve according to a particular first order linear ODE system when ω is fixed. Lemma VI.23 makes this precise when this time is positive. We leave the investigation and interpretation of the case when it is zero for future work. Also, note that our agents effectively assume that this time is T almost surely.

Lemma VI.23. *Suppose that $\det A$ has a root on $[0, T]$. If $t_e > 0$, then $S^{exc}(\omega)$, the $X_j^{\theta_j^\star}(\omega)$'s and the $\theta_j^\star(\omega)$'s are all uniquely defined and continuous on $[0, t_e)$. More-*

over, letting the u-superscript signify restriction to the uncertain agents,[12] *the uncertain agents' strategies are characterized by*

$$A(t)\,\theta_t^{u,\star}(\omega) = B(t)\,X_t^{u,\theta^\star}(\omega) + C(t,\omega)\,, \qquad t \in (0,t_e)$$
$$X_0^{u,\theta^\star}(\omega) = x^u. \tag{6.14}$$

When $\det A$ *does not have a root on* $[0,T]$*, the same statements hold after replacing* t_e *with* T*.*

Proof. See Subsection 6.9.1. □

Lemma VI.23 does not address the behavior of our uncertain agents' inventories and trading rates as $t \uparrow t_e$ or $t \uparrow T$. The difficulties are that A is non-invertible at t_e, while B's entries explode at T (see Lemma VI.21).

The approach for resolving these issues is well-established (see Chapter 6 of [87]). We sketch the key points when $\det A$ has a root on $[0,T]$ and $t_e > 0$. Analyzing the effects of B's explosion at T is similar (see Lemma VI.28).

We begin by considering the homogeneous equation corresponding to (6.14):

$$A(t)\,\dot{X}_t^u(\omega) = B(t)\,X_t^u(\omega)\,, \qquad t \in (0,t_e)$$
$$X_0^u(\omega) = x^u. \tag{6.15}$$

We change notation to emphasize that (6.15) no longer describes the uncertain agents' optimal strategies. We next write (6.15) in a more convenient form.

Lemma VI.24. *Suppose that* $\det A$ *has a root on* $[0,T]$ *and* $t_e > 0$*. Near* t_e*, the solution of (6.15) satisfies*

$$(t-t_e)^{m+1}\,\dot{X}_t^u(\omega) = D(t)\,X_t^u(\omega)\,. \tag{6.16}$$

[12]For instance, $\theta_t^{u,\star}(\omega)$ denotes the first K-entries of $\theta_t^\star(\omega)$.

In (6.16), m is a nonnegative integer such that the multiplicity of the zero of $\det A$ *at t_e is $(m+1)$. D is a particular analytic map for which $D(t_e)$ has rank 0 or 1 (see (6.47)).*

Proof. See Subsection 6.9.2. □

Definition VI.25. If $\det A$ has a root on $[0,T]$ and $t_e > 0$, we let m, D, and f be defined as in Lemma VI.24's proof (see (6.46) and (6.47)). Also, $D(t_e)$ has at most one non-zero eigenvalue (see Lemma VI.24's proof), which we denote by λ.

Remark VI.26. Unless $D(t_e) = 0$, $D(t_e)$ has rank 1 (see Lemma VI.24's proof). Hence, we can find v, $\hat{v} \in \mathbb{R}^K$ such that

$$v\hat{v}^\top = D(t_e) \quad \text{and} \quad \hat{v}^\top v = \lambda.$$

Moreover, v is an eigenvector of $D(t_e)$ corresponding to λ. While v and $\hat{v}$ are not unique, algorithms are available to compute an example of such a pair ([200]). In future work, we may use this decomposition to investigate the occurrence of mini-flash crashes in broader cases than those considered in Section 6.7.

Suppose that $D(t_e) \neq 0$. Since $t_e < T$, the coefficients of (6.15) are analytic in a neighborhood of t_e (see Lemma VI.21). It follows that (6.15) has a *singular point of the first kind* at t_e when $m = 0$ in Lemma VI.24 (see Chapter 6 of ([87])). Otherwise, the singular point is of the *second kind*.[13] In the former case, the fundamental solution of (6.15) near t_e is the product of a certain analytic function with a matrix exponential.

The analysis of solution behavior when there is a singular point of the second kind at t_e is significantly more difficult. For instance, while we may be able to find a formal series solution for (6.15) near t_e, it may converge at just one point.[14] We do

[13] We adopt the nomenclature from Coddington & Carlson ([87]); however, other sources refer to such points as *regular* and *irregular singular points*, respectively ([137]). There are nonequivalent definitions of these terms too.

[14] See the books by Wasow ([233]) and Ilyashenko & Yakovenko ([137]) for detailed discussions on these issues.

not consider scenarios with such singularities in the present work, as our examples in Section 6.7 do not exhibit them (see Lemma VI.34).

As soon as we have the fundamental solution near t_e, we use variation of parameters to solve (6.14). This gives our uncertain agents' optimal inventories. We immediately get their optimal trading rates by differentiating and the corresponding execution price by plugging all agents' strategies into (6.13).

This discussion is made precise in the next result.

Lemma VI.27. *Suppose that* $\det A$ *has a root on* $[0,T]$*,* $t_e > 0$*, and* $m = 0$*. If* $\lambda \notin \mathbb{Z}$*,*[15] *then for some small* $\rho > 0$*,*

$$X_t^{u,\theta^\star}(\omega) = P(t)\left[\sum_{j=1}^{K-1}\left(y_j(\omega) - \int_{t_e-\rho}^{t}\frac{F_j(s,\omega)}{|s-t_e|}\,ds\right)v_j + |t-t_e|^\lambda\left(y_K(\omega) - \int_{t_e-\rho}^{t}\frac{F_K(s,\omega)}{|s-t_e|^{1+\lambda}}\,ds\right)v_K\right] \tag{6.17}$$

for $t \in (t_e - \rho, t_e)$*. Here,*

- $\{v_1, \ldots, v_K\}$ *is an eigenbasis for* $D(t_e)$ *(*v_K *corresponds to* λ*);*
- P *is a (non-singular-)matrix-valued analytic function on* $[t_e - \rho, t_e]$ *such that* $P(t_e) = I_K$ *(see (6.48));*
- $\{y_1(\omega), \ldots, y_K(\omega)\}$ *are constants (see (6.51));*
- *and* $\{F_1(\cdot,\omega), \ldots, F_K(\cdot,\omega)\}$ *are continuous real-valued functions on* $[t_e - \rho, t_e]$ *(see (6.51)).*

[15] In Section 6.7, we can always slightly perturb our parameters, if necessary, to ensure that $\lambda \notin \mathbb{Z}$ (see Lemma VI.37).

Generally, the $\lambda \in \mathbb{Z}$ case may or may not be more difficult to avoid. We leave this point for future work. In principle, there could be three additional scenarios to consider: $D(t_e) = 0$, $D(t_e)$ is a non-zero nilpotent matrix, and $\lambda \neq 0$.

When $D(t_e) = 0$, the matrix exponential in the fundamental solution of (6.15) can be dropped (see Sections 2.3 and 5.6 of [87]). If $D(t_e)$ is a non-zero nilpotent matrix, the series representation of the matrix exponential terminates after $(K-1)$ terms, maybe fewer, as the degree of $D(t_e)$ is no higher than K.

In the last case, the argument is less transparent. A crucial recursion used in the determination of P is no longer valid, necessitating an intricate change of variables (see Chapter 6 of [87]). Consequently, the fundamental solution of (6.15) is that in (6.48) but with $D(t_e)$ replaced by a more opaque matrix, which significantly complicates further analysis.

We get $\theta^{u,\star}(\omega)$ *and* $S^{exc}(\omega)$ *on* $(t_e-\rho,t_e)$ *by differentiating (6.17) and by substituting* $X^{\theta^\star}(\omega)$ *and* $\theta^\star(\omega)$ *into (6.13), respectively.*[16]

Proof. See Subsection 6.9.3. □

Lemma VI.28. *Suppose that* $\det A$ *does not have a root on* $[0,T]$. *Then* $S^{exc}(\omega)$, *the* $X_j^{\theta_j^\star}(\omega)$*'s and the* $\theta_j^\star(\omega)$*'s are all uniquely defined and continuous on* $[0,T]$. *Moreover,*

$$\lim_{t\uparrow T} X_t^{\theta^\star}(\omega)=0. \tag{6.18}$$

Remark VI.29. Each agent believes that his terminal inventory will be zero almost surely (see (6.5)). Lemma VI.28 specifies conditions under which the agents are effectively correct in this regard.

Proof. See Subsection 6.9.4. □

Remark VI.30. For general parameter choices, using Lemma VI.27 to investigate the occurrence of mini-flash crashes may be difficult. Here are the key challenges:

a) Transparent conditions governing the existence of a root of $\det A$ on $[0,T]$ are not immediate.[17]

b) It is not yet obvious when, if ever, the multiplicity of $\det A$'s root at t_e will be 1.

c) It is unclear that we can ensure that $\lambda\notin\mathbb{Z}$, even after a perturbation of our parameters.

d) More analysis of (6.17) is needed to characterize potential explosions in the coordinates of $X_t^{u,\theta^\star}$ and $\theta_t^{u,\star}$ as $t\uparrow t_e$.

[16] Recall that the certain agents' inventories and trading rates were found in Corollary VI.15.

[17] Once we have such conditions, easily checking whether or not $t_e>0$ would be presumably trivial but could be troublesome as well.

e) Determining whether or not S^{exc} explodes requires a more thorough study of (6.13) and (6.17).

Resolving (a) and (b) is rather intractable, unless K is small or our uncertain agents are fairly similar (see Definition VI.19). Completing the studies suggested by (c) and (d) requires further knowledge of λ and the eigenbasis $\{v_1, \dots, v_K\}$ of $D(t_e)$. Even then, the y_j's and the F_j's in (6.51) may be quite opaque and pose obstacles. These observations further restrict the size of K or the differences among our agents. After all of these restrictions, finishing (e) may still not be straightforward, as in principle, the coordinates of $X_t^{u,\theta^\star}$ or $\theta_t^{u,\star}$ might explode at the same rates but in opposite directions.

Hence, we investigate mini-flash crashes only in the context of a particularly tractable class of examples (see Section 6.7). We offer a rough summary of our mathematical findings in Theorem VI.31; however, the details and practical connections to mini-flash crashes are in Section 6.7.

We leave the study of other scenarios for future work. For instance, it would be interesting to know whether or not we could observe unbounded price oscillations near t_e or mini-flash crashes in the absence of synchronized trading (see Section 6.3).

Theorem VI.31. *When our agents are as characterized in Section 6.4 but the risky asset's price evolves as in Section 6.5, at least three cases emerge (see Lemmas VI.28, VI.34, VI.39, and VI.42 for precise statements): There are broad sufficient conditions on our deterministic parameters such that*

i) S^{exc}*, the* $X_j^{\theta_j^\star}$*'s and the* $\theta_j^\star$*'s are all uniquely defined and continuous on* $[0,T]$ *and*

$$\lim_{t \uparrow T} X_t^{\theta^\star} = 0 \quad \tilde{P} - a.s.;$$

ii) S^{exc}*, the* $X_j^{u,\theta_j^\star}$*'s, and the* $\theta_j^{u,\star}$*'s explode as* $t \uparrow t_e$ $\tilde{P}$*-a.s.;*

iii) or all coordinates of $X^{\theta^\star}$ *have a finite limit but* S^{exc} *and the* $\theta_j^{u,\star}$*'s explode as* $t \uparrow t_e$ $\tilde{P}$*-a.s.*

In scenarios (ii) and (iii), all explosions occur in the same random direction: $+\infty$ *or* $-\infty$*. The* $\tilde{P}$*-probability of infinite spikes (crashes) tends to either 0 or 1 as* $t \uparrow t_e$*; however, it is positive for any fixed* $t < t_e$*.*

While our conditions are deterministic, no agent knows the critical parameters in these calculations. In particular, our agents believe that a mini-flash crash is a null event.

Proof. The result follows immediately from Lemmas VI.28, VI.34, VI.37, VI.39, and VI.42. □

6.7 Semi-Symmetric Uncertain Agents

In Section 6.7, we thoroughly analyze a broad but tractable class of scenarios. This will enable us to both theoretically and numerically investigate the occurrence of mini-flash crashes.

Based on Remark VI.30, we specify that our uncertain agents' parameters are identical, except for their initial inventories x_j, means of their initial drift priors μ_j, and their initial estimates for the fundamental price $S_{j,0}$. Such agents are nearly symmetric, so we call them *semi-symmetric*.

Definition VI.32. We say that our uncertain agents are *semi-symmetric* when there are positive constants

$$\tilde{\eta}_{tem}, \quad \eta_{tem}, \quad \tilde{\eta}_{per}, \quad \eta_{per}, \quad \nu^2, \quad \text{and} \quad \kappa$$

such that for each $i \in \{1, \ldots, K\}$

$$\tilde{\eta}_{i,tem} = \tilde{\eta}_{tem}, \quad \eta_{i,tem} = \eta_{tem}, \quad \tilde{\eta}_{per} = \tilde{\eta}_{i,per},$$
$$\eta_{i,per} = \eta_{per}, \quad \nu_i^2 = \nu^2, \qquad \kappa_i = \kappa.$$

Definition VI.32 implies that the diagonal entries of A are identical, as are the off-diagonal entries. The same is true for B (see Definition VI.19). Such a simplification considerably reduces the difficulties in computing $\det A$, λ, and an eigenbasis for $D(t_e)$ (see (6.61) and Lemma VI.37). The x_j's, μ_j's, and $S_{j,0}$'s only enter in C, which also has a nice structure (see (6.71)).

For the rest of Section 6.7, we assume that our uncertain agents are semi-symmetric but place no restrictions on the certain agents. Our theoretical results are contained in Subsection 6.7.1. In Subsections 6.7.2 - 6.7.4, we provide figures for key conclusions on mini-flash crashes (see Subsection 6.2.1). With these plots, our goal is to highlight the features of our model, not to recreate any specific historical scenario.

6.7.1 Results

Notation VI.33. If our uncertain agents are semi-symmetric, the τ_j's and the Φ_j's are the same for $j \leq K$ (see Definitions VI.7, VI.19, and VI.32). We denote these functions by τ and Φ, respectively.

Lemma VI.34. *Suppose that the uncertain agents are semi-symmetric. Then* $\det A$ *has a root on* $[0, T]$ *and* $t_e > 0$ *if and only if*

$$(K\tilde{\eta}_{tem} - \eta_{tem})\,\Phi(0) > 2. \tag{6.19}$$

In this case, the zero of $\det A(\cdot)$ *at* t_e *is of multiplicity 1.*

Remark VI.35. Definitions VI.7 and VI.19 enable us to re-write (6.19) as

$$\frac{\nu^2 (K\tilde{\eta}_{tem} - \eta_{tem}) \tanh\left(\dfrac{T}{2}\sqrt{\dfrac{\kappa}{\eta_{tem}}}\right)}{\sqrt{\eta_{tem}\kappa}} > 2. \tag{6.20}$$

By varying our parameters in (6.20) one at a time, (6.19) can be interpreted as discussed in Subsection 6.2.1:

i) (6.19) holds when ν^2 is high. Since ν^2 is the variance of the uncertain agents' drift priors, we are led to (f) in Subsection 6.2.1.

ii) (6.19) holds when $(K\tilde{\eta}_{tem} - \eta_{tem})$ is high. A given uncertain agent believes that his own temporary impact parameter is η_{tem}, while the actual collective temporary impact parameter induced by the uncertain agents is $K\tilde{\eta}_{tem}$. Then $(K\tilde{\eta}_{tem} - \eta_{tem})$ is large whenever each uncertain agent severely underestimates his own temporary impact or there are many uncertain agents, giving (g) in Subsection 6.2.1.

iii) (6.19) holds when T is high. Since $[0, T]$ is our time horizon, we get (d) in Subsection 6.2.1. Note that T must be small enough for our agents' modeling rationale to hold (see Section 6.4); however, T need not be too large here, as the value of tanh reaches 95% of its supremum on $[0, \infty)$ for arguments greater than 1.8.

iv) (6.19) holds when κ is low. We conclude (e) in Subsection 6.2.1, as κ measures our uncertain agents' aversion to volatility risks (see Subsection 6.4.4). Observe that both the numerator and the denominator of the LHS in (6.19) roughly look like $\sqrt{\kappa}$ for small κ; however, when κ is large, the whole LHS looks like $1/\sqrt{\kappa}$ since tanh is bounded by 1 on $[0, \infty)$.

Proof. See Subsection 6.10.1. □

Remark VI.36. As observed in (6.62), when $\det A$ has a root on $[0, T]$, we have

$$\Phi\left(t_e\right) = \frac{2}{K\tilde{\eta}_{tem} - \eta_{tem}}. \tag{6.21}$$

No agent would think to compute t_e since they all believe that a mini-flash crash is a null event; however, (6.21) makes it especially clear that they could not do so anyway.

Lemma VI.37. *Suppose that the uncertain agents are semi-symmetric and (6.19) holds. Then*

$$\lambda = \frac{2\left[\sqrt{\frac{\kappa}{\eta_{tem}}}\coth\left(\tau\left(t_e\right)\right) - 2\left(\frac{K\tilde{\eta}_{per} - \eta_{per}}{K\tilde{\eta}_{tem} - \eta_{tem}}\right)\right]}{\left(K\tilde{\eta}_{tem} - \eta_{tem}\right)\dot{\Phi}\left(t_e\right)} \tag{6.22}$$

and the corresponding eigenvector is $v_K = [1, \ldots, 1]^\top$*. By slightly perturbing* $\tilde{\eta}_{per}$ *and/or* η_{per}*, if necessary, we can ensure that* $\lambda \notin \mathbb{Z}$*. In this case,* $D\left(t_e\right)$ *is diagonalizable and the remaining vectors in an eigenbasis for* $D\left(t_e\right)$ *(all with the eigenvalue zero) are*

$$v_1 = [-1, 1, 0, \ldots, 0]^\top \quad , \ldots , \quad v_{K-1} = [-1, 0, \ldots, 0, 1]^\top .$$

Remark VI.38. With the exceptions of $\tilde{\eta}_{per}$ and η_{per}, all parameters in (6.22) determine whether or not $\det A$ has a root on $[0, T]$ (see Lemma VI.34). They also fix the value of t_e (see Remark VI.36). Hence, to interpret (6.22), we only consider the roles of $\tilde{\eta}_{per}$ and η_{per}. These parameters enter (6.22) via

$$\frac{K\tilde{\eta}_{per} - \eta_{per}}{K\tilde{\eta}_{tem} - \eta_{tem}}. \tag{6.23}$$

Intuitively, (6.23) can be viewed as the ratio of two terms: The numerator measures how far a given uncertain agent's estimate of his own permanent impact is from the uncertain agents' actual collective permanent impact. The denominator, which must be positive due to Lemma VI.34, is the corresponding measure for the temporary impact. One might call (6.23) a *mistake ratio.*

Since $\dot{\Phi}\left(t_e\right) < 0$ by Lemma VI.21, λ is positive only when (6.23) is high enough. We have $\lambda < 0$ when the uncertain agents' total permanent impact and a single uncertain agent's' estimate of his own permanent impact are too close or when his estimate exceeds the cumulative permanent impact. More precisely,

$$\begin{aligned} \{\lambda > 0\} \quad &\Longleftrightarrow \quad \left\{ \frac{1}{2}\sqrt{\frac{\kappa}{\eta_{tem}}} \coth\left(\tau\left(t_e\right)\right)\left(K\tilde{\eta}_{tem} - \eta_{tem}\right) < K\tilde{\eta}_{per} - \eta_{per} \right\} \\ \{\lambda < 0\} \quad &\Longleftrightarrow \quad \left\{ \frac{1}{2}\sqrt{\frac{\kappa}{\eta_{tem}}} \coth\left(\tau\left(t_e\right)\right)\left(K\tilde{\eta}_{tem} - \eta_{tem}\right) > K\tilde{\eta}_{per} - \eta_{per} \right\}. \end{aligned} \tag{6.24}$$

Whether a mini-flash crash is accompanied by high or low trading volumes is effectively determined by which inequality in (6.24) holds (see Lemmas VI.39 and VI.42 and Subsections 6.2.1, 6.7.3, and 6.7.4).

Proof. See Subsection 6.10.2. □

Lemma VI.39. *Suppose that the uncertain agents are semi-symmetric and (6.19) holds. Assume that* $\lambda \notin \mathbb{Z}$ *and* $\lambda < 0$ *(see Lemma VI.37). Let* ρ, $y_K(\omega)$, *and* $F_K(\cdot, \omega)$ *be defined as in Lemma VI.27. Then*

$$\begin{aligned} &\left\{ y_K(\omega) > \lim_{t \uparrow t_e} \int_{t_e - \rho}^{t} \frac{F_K(s,\omega)}{|s - t_e|^{1+\lambda}}\, ds \right\} \\ &\quad \Longrightarrow \quad \left\{ \lim_{t \uparrow t_e} X_t^{u,\theta^\star}(\omega) = \lim_{t \uparrow t_e} \theta_t^{u,\star}(\omega) = [+\infty, \ldots, +\infty]^\top, \quad \lim_{t \uparrow t_e} S_t^{exc}(\omega) = +\infty \right\} \end{aligned} \tag{6.25}$$

and

$$\begin{aligned} &\left\{ y_K(\omega) < \lim_{t \uparrow t_e} \int_{t_e - \rho}^{t} \frac{F_K(s,\omega)}{|s - t_e|^{1+\lambda}}\, ds \right\} \\ &\quad \Longrightarrow \quad \left\{ \lim_{t \uparrow t_e} X_t^{u,\theta^\star}(\omega) = \lim_{t \uparrow t_e} \theta_t^{u,\star}(\omega) = [-\infty, \ldots, -\infty]^\top, \quad \lim_{t \uparrow t_e} S_t^{exc}(\omega) = -\infty \right\}. \end{aligned} \tag{6.26}$$

Moreover,

i) The integral limits in (6.25) and (6.26) exist and are finite.

ii) Either (6.25) or (6.26) holds $\tilde{P}$*-a.s.*

iii) At $t_e - \rho$, the events (6.25) and (6.26) both have positive $\tilde{P}$-probability; however, the $\tilde{P}$-probability of one event tends to 1 (while the other tends to 0) if we let $\rho \downarrow 0$.

Remark VI.40. Although we fixed ω in Notation VI.18, by abuse, we view it as varying for our probabilistic statements in Lemmas VI.39 and VI.42.

Remark VI.41. Since $P(t_e) = I_K$ (see Lemma VI.27), (6.51) and Lemma VI.37 imply that $y_K(\omega)$ will be large and positive when the uncertain agents hold significant, similar long positions. $y_K(\omega)$ will be of high magnitude but negative, if the uncertain agents carry substantial, similarly-sized short positions. Hence, a spike in $S^{exc}(\omega)$ is more likely when the uncertain agents are synchronized aggressive buyers, while the odds of a collapse improve when they are synchronized heavy sellers. These effects play the deciding role as $t \uparrow t_e$, as the integral limits in (6.25) and (6.26) are finite.

Still, due to how we can decompose F_K in our case (see (6.74)), large fluctuations in the fundamental price can make the mini-flash crash's direction unclear until just before t_e (see Figure 6.9).

Proof. See Subsection 6.10.3. □

Lemma VI.42. *Suppose that the uncertain agents are semi-symmetric and (6.19) holds. Assume that $\lambda \notin \mathbb{Z}$ and $\lambda > 0$ (see Lemma VI.37). Then $\tilde{P}$-a.s.,*

$$\lim_{t \uparrow t_e} X_t^{\theta^\star}(\omega)$$

exists in $\mathbb{R}^N$. If any coordinates of $\theta_t^{u,\star}(\omega)$ explode, then $S_t^{exc}(\omega)$ and all coordinates of $\theta_t^{u,\star}(\omega)$ explode in the same direction. For instance, when $\lambda > 1$,

$$\left\{\lim_{t \uparrow t_e}\left[|t - t_e|^{\lambda - 1}\int_{t_e - \rho}^{t}\frac{\tilde{W}_s(\omega) - \tilde{W}_t(\omega)}{|s - t_e|^{1+\lambda}}\,ds\right] = +\infty\right\} \tag{6.27}$$

$$\Longrightarrow \quad \left\{\lim_{t \uparrow t_e}\theta_t^{u,\star}(\omega) = [+\infty, \ldots, +\infty]^\top, \quad \lim_{t \uparrow t_e} S_t^{exc}(\omega) = +\infty\right\}$$

and

$$\left\{\lim_{t\uparrow t_e}\left[\left|t-t_e\right|^{\lambda-1}\int_{t_e-\rho}^{t}\frac{\tilde{W}_s\left(\omega\right)-\tilde{W}_t\left(\omega\right)}{\left|s-t_e\right|^{1+\lambda}}\,ds\right]=-\infty\right\} \tag{6.28}$$

$$\Longrightarrow \quad \left\{\lim_{t\uparrow t_e}\theta_t^{u,\star}\left(\omega\right)=\left[-\infty,\ldots,-\infty\right]^{\top},\quad \lim_{t\uparrow t_e}S_t^{exc}\left(\omega\right)=-\infty\right\}.$$

Moreover,

i) Either (6.27) or (6.28) holds $\tilde{P}$-a.s.

ii) At $t_e-\rho$, the events (6.27) and (6.28) both have positive $\tilde{P}$-probability; however, the $\tilde{P}$-probability of one event tends to 1 (while the other tends to 0) if we let $\rho\downarrow 0$.

Remark VI.43. We make no rigorous statement regarding the $\lambda\in(0,1)$ case. Most of Lemma VI.42's proof would still be valid (see Subsection 6.10.3); however, the final estimates are especially convenient when $\lambda>1$ (see (6.85) - (6.89)). The overarching purpose of Lemma VI.42 is only to illustrate that mini-flash crashes can occur in low trading volume environments (see Subsection 6.2.1). Nevertheless, we suspect that mini-flash crashes might unfold when $\lambda\in(0,1)$, e.g., see Subsection 6.7.3 and (6.85) - (6.89).

Proof. See Subsection 6.10.3. □

6.7.2 Example 1: No mini-flash crash

Our mini-flash crashes do not always occur (see Lemmas VI.28 and VI.34). In Subsection 6.7.2, we illustrate this by numerically simulating a scenario in which $\det A$ has no root on $[0,T]$.

By Lemma VI.34 and (6.61), we know that $\det A$ is non-vanishing on $[0,T]$ if and only if

$$\left(K\tilde{\eta}_{tem}-\eta_{tem}\right)\Phi\left(0\right)<2. \tag{6.29}$$

One selection of parameters for which (6.29) is satisfied is

$$\begin{array}{llll} N = 3, & K = 2, & T = 1, & S_0 = 100, \\ \tilde{\beta} = 1, & \tilde{\eta}_{tem} = 1, & \eta_{tem} = 0.75, & \tilde{\eta}_{per} = 1, \\ \eta_{per} = 1, & \nu^2 = 2, & \kappa = 5, & x_1 = 2, \\ x_2 = -2, & \mu_1 = 15, & \mu_2 = -10, & S_{1,0} = 100, \\ S_{2,0} = 100, & \tilde{\eta}_{3,tem} = 1, & \eta_{3,tem} = 1, & \tilde{\eta}_{3,per} = 1, \\ \mu_3 = -3, & \nu_3^2 = 2, & \kappa_3 = 5, & x_3 = 2. \end{array} \tag{6.30}$$

In fact, the LHS of (6.29) then equals 1.1095. Observe that there is no need to specify $\eta_{3,per}$ and $S_{3,0}$ as they are irrelevant (see Corollary VI.15, Definition VI.19, and Lemma VI.23). Again, our purposes are only illustrative here, and we leave the reproduction of a specific practically meaningful scenario for a future work.

Since $K = 2$ and $N = 3$, we have two uncertain agents and one certain agent in the coming plots. We label the corresponding curves with $U1$, $U2$, and $C1$. For example, the label $U1$ will signify a quantity for Agent 1, the first uncertain agent. In Figures 6.1 and 6.2, we plot inventories and trading rates. The execution price is depicted in Figure 6.3.

The diagrams exhibit all of the important qualities that we expect based upon our theoretical results. Here are a few key features:

i) All agents liquidate their positions by the terminal time T (see (6.18) and Figure 6.1).

ii) $S^{exc}(\omega)$, the $X_j^{\theta_j^\star}(\omega)$'s and the $\theta_j^\star(\omega)$'s are all continuous on $[0, T]$ (see Lemma VI.28 and Figures 6.1 - 6.3).

iii) The uncertain agents' trading rates appear to exhibit a Brownian component (see Lemma VI.9 and Figure 6.2).

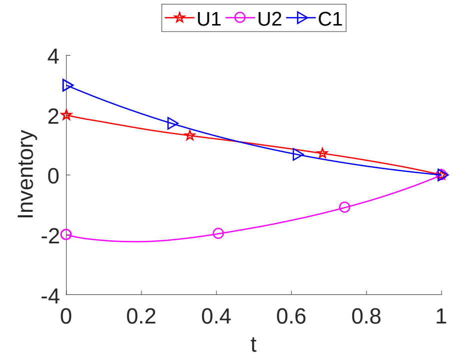

Figure 6.1: Depiction of the agents' inventories in Subsection 6.7.2.

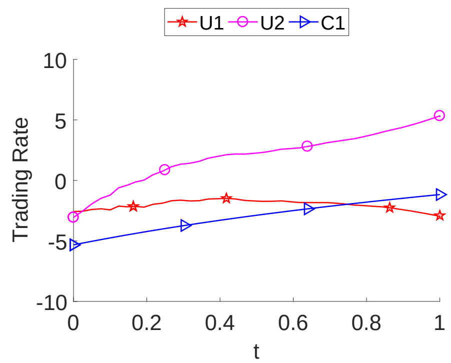

Figure 6.2: Depiction of the agents' trading rates in Subsection 6.7.2.

iv) The certain agent's trading rate appears to be smooth on $[0, T]$ (see Corollary VI.15 and Figure 6.2).

v) The agents need not either strictly buy or strictly sell throughout $[0, T]$ (see Subsection 6.4.3, Lemma VI.9 and Figure 6.2).

vi) Even so, the agents may decide to strictly buy or strictly sell throughout $[0, T]$ (see Subsection 6.4.3, Lemma VI.9 and Figure 6.2).

vii) The uncertain agents' trading rates do not appear to synchronize (see Figure 6.2).

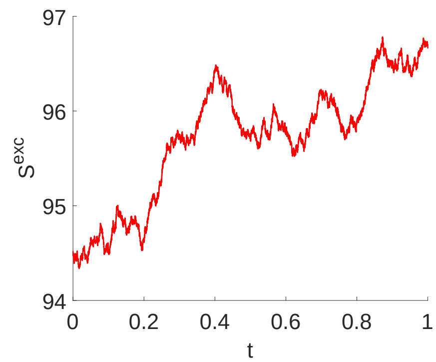

Figure 6.3: Depiction of the execution price in Subsection 6.7.2.

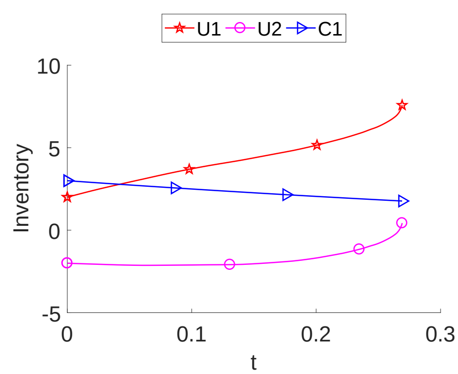

Figure 6.4: Depiction of the agents' inventories in Subsection 6.7.3.

6.7.3 Example 2: A mini-flash crash with low trading volume

Our mini-flash crashes can be accompanied by low trading volumes (see Lemma VI.42). In Subsection 6.7.3, we visualize this by studying a concrete scenario in which $\det A$ has a root on $[0,T]$; $t_e > 0$; the zero of $\det A$ at t_e is of multiplicity 1; $\lambda \notin \mathbb{Z}$; and $\lambda > 0$. The behavior of the $X_j^{\theta_j^\star}(\omega)$'s is then characterized by Corollary VI.15 and Lemma VI.42. Lemma VI.42 would rigorously describe $S_t^{exc}(\omega)$ and the $\theta_{j,t}^\star(\omega)$'s as $t \uparrow t_e$, if $\lambda > 1$. To improve the quality of our plots, we consider a situation where $\lambda \in (0,1)$ instead (see Remark VI.43).

By Lemmas VI.34 and VI.37, we must select parameters such that (6.19) is satisfied and

$$\lambda = \frac{2\left[\sqrt{\frac{\kappa}{\eta_{tem}}}\coth\left(\tau\left(t_e\right)\right) - 2\left(\frac{K\tilde{\eta}_{per} - \eta_{per}}{K\tilde{\eta}_{tem} - \eta_{tem}}\right)\right]}{\left(K\tilde{\eta}_{tem} - \eta_{tem}\right)\dot{\Phi}\left(t_e\right)} \tag{6.31}$$

is a positive non-integer. We can keep most of our choices in (6.30) the same and only make a few revisions:

$$\begin{aligned} &\tilde{\eta}_{tem} = 0.5, \quad &&\eta_{tem} = 0.2, \quad &&\tilde{\eta}_{per} = 0.8, \\ &\eta_{per} = 0.025, \quad &&\nu^2 = 3, \quad &&\kappa = 1. \end{aligned} \tag{6.32}$$

As in Subsection 6.7.2, we do not seek to replicate a particular historical situation. We immediately get (6.19), as its LHS is 4.3302. Using Remark VI.36 and (6.31), we can show that

$$t_e = 0.2691 \quad \text{and} \quad \lambda = 0.5939.$$

Again, we have two uncertain agents and one certain agent. We retain the $\{U1, U2, C1\}$- labeling system from Subsection 6.7.2. The inventories, trading rates, and execution price are plotted in Figures 6.4 - 6.6. To aid our illustration, we truncate the time domains in Figures 6.5 - 6.6 to

$$\left[0, 0.75\left(t_e - 10^{-6}\right)\right] \quad \text{and} \quad \left[0, t_e - 10^{-6}\right]$$

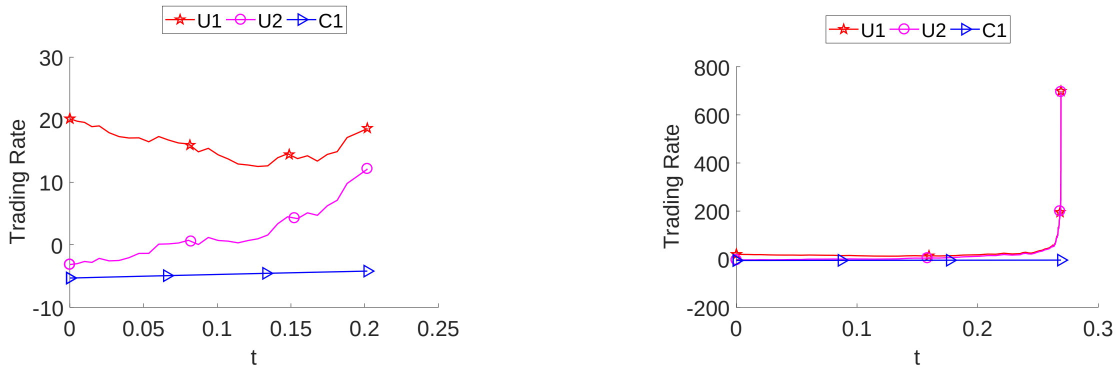

Figure 6.5: Depiction of the agents' trading rates in Subsection 6.7.3.

for the left and right plots, respectively.

The qualities that we expect based upon Corollary VI.16, Lemma VI.42, and Remark VI.43 are all present. We offered some applicable comments in Subsection 6.7.2, so we only add a few new observations here.

i) All agents' inventories approach a finite limit as $t \uparrow t_e$ (see Lemma VI.42 and Figure 6.4).

ii) The execution price and the uncertain agents' trading rates explode as $t \uparrow t_e$ (see Lemma VI.42, Remark VI.43 and Figures 6.5 - 6.6).

iii) The uncertain agents' trading rates synchronize as $t \uparrow t_e$ (see Lemma VI.42, Remark VI.43, and Figure 6.5).

iv) That an explosion in $S^{exc}(\omega)$ will occur as well as its direction becomes increasingly obvious as $t \uparrow t_e$; however, it is not necessarily clear at first (see Lemma VI.42, Remark VI.43, and Figure 6.6).

6.7.4 Example 3: A mini-flash crash with high trading volume

Our mini-flash crashes can also be accompanied by high trading volumes (see Lemma VI.39). We illustrate this in Subsection 6.7.4 by simulating a case in which

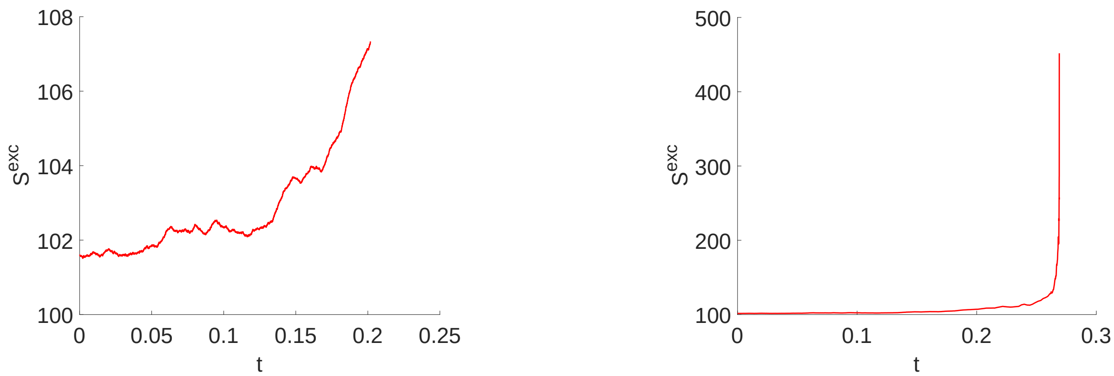

Figure 6.6: Depiction of the execution price in Subsection 6.7.3.

$\det A$ has a root on $[0,T]$; $t_e > 0$; the zero of $\det A$ at t_e is of multiplicity 1; $\lambda \notin \mathbb{Z}$; and $\lambda < 0$. The behaviors of $S^{exc}(\omega)$, the $X_j^{\theta_j^\star}(\omega)$'s, and the $\theta_j^\star(\omega)$'s are then described by Corollary VI.15 and Lemma VI.39.

We especially wish to emphasize the stochastic explosion direction and do this in two ways.

First, we choose the same deterministic parameters to create Figures 6.7 - 6.12. The difference is that one realization of $\tilde{W}_\cdot$ is used in Figures 6.7 - 6.9, while another is used in Figures 6.10 - 6.12. We denote the corresponding ω's by ω_{up} and ω_{dn}, since there are spikes and crashes in the former and latter plots, respectively.

Second, Figures 6.7 - 6.9 themselves suggest that the explosion direction is random. This is particularly true in Figures 6.8 - 6.9, since we initially notice that the price rapidly rises as the uncertain agents' buying rates synchronize. Only moments before the mini-flash crash do we see the price collapsing and the uncertain agents' aggressively selling together.

Now, we need to choose parameters such that (6.19) is satisfied and

$$\lambda = \frac{2\left[\sqrt{\frac{\kappa}{\eta_{tem}}}\coth\left(\tau\left(t_e\right)\right) - 2\left(\frac{K\tilde{\eta}_{per} - \eta_{per}}{K\tilde{\eta}_{tem} - \eta_{tem}}\right)\right]}{\left(K\tilde{\eta}_{tem} - \eta_{tem}\right)\dot{\Phi}\left(t_e\right)}$$

is a negative non-integer due to Lemmas VI.34 and VI.39. Compared to Subsection

6.7.3, we set

$$\tilde{\eta}_{per} = 0.5, \quad \eta_{per} = 0.5$$

and keep every other parameter the same. As in Subsections 6.7.2 - 6.7.3, we do not have in mind a special historical example here. Since we have only changed $\tilde{\eta}_{per}$ and η_{per}, the values of $(K\tilde{\eta}_{tem} - \eta_{tem})\,\Phi(0)$ and t_e do not differ from Subsection 6.7.3; however, λ is now negative:

$$(K\tilde{\eta}_{tem} - \eta_{tem})\,\Phi(0) = 4.3302, \quad t_e = 0.2691, \quad \text{and} \quad \lambda = -0.4531.$$

The numbers of uncertain and certain agents are still two and one, respectively. We also retain the $\{U1, U2, C1\}$-labeling system from Subsections 6.7.2 - 6.7.3. Figures 6.7 and 6.10 depict the agents' inventories. We plot the agents' trading rates in Figures 6.8 and 6.11. The execution price appears in Figures 6.9 and 6.12. To help with our visualization, the time domains in the left plots in Figures 6.7 - 6.9 and Figures 6.10 - 6.12 are truncated to $[0, 0.94\,(t_e - 10^{-6})]$ and $[0, 0.75\,(t_e - 10^{-6})]$, respectively.

Our observations regarding Figures 6.7 - 6.12 are in agreement with Corollary VI.15 and Lemma VI.39. We have already made note of many important aspects in Subsections 6.7.2 - 6.7.3 and only remark upon the new details.

i) The execution price, as well as the uncertain agents' inventories and trading rates, all explode in the same direction as $t \uparrow t_e$ (see Lemma VI.39 and Figures 6.7 - 6.12).

ii) The explosions take place at the deterministic time t_e (see Lemma VI.39 and Figures 6.7 - 6.12).

iii) The explosion direction depends on $\omega \in \tilde{\Omega}$ (see Lemma VI.39 and Figures 6.7 - 6.12).

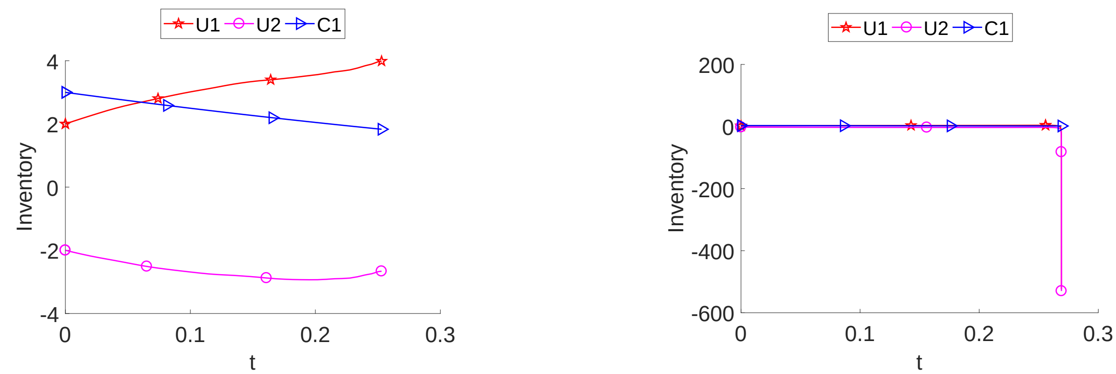

Figure 6.7: Depiction of the agents' inventories for ω_{dn} in Subsection 6.7.4.

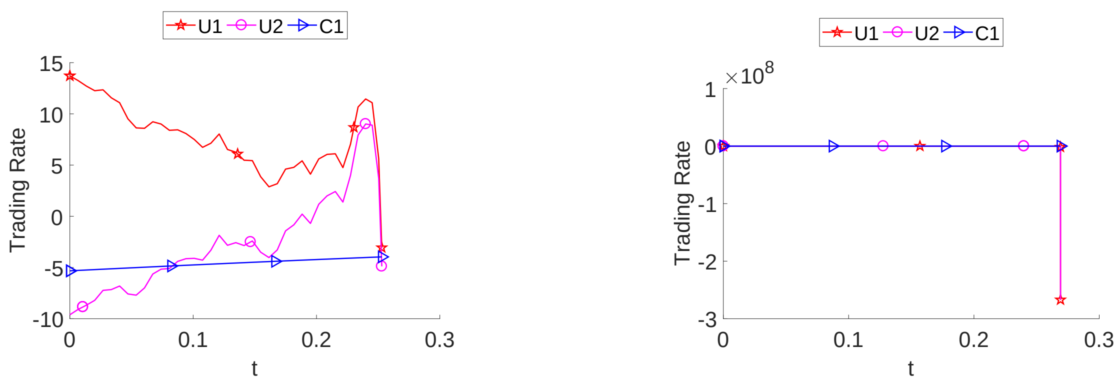

Figure 6.8: Depiction of the agents' trading rates for ω_{dn} in Subsection 6.7.4.

iv) The explosion direction cannot be known with complete certainty before t_e (see Lemma VI.39 and Figures 6.7 - 6.12).

v) The explosion rates in the price and uncertain agents' trading rates in Subsection 6.7.3 are slower than in Subsection 6.7.4 (see Figures 6.5 - 6.6, Figures 6.8 - 6.9, and Figures 6.11 - 6.12). We did not explicitly state this previously; however, this is to be expected since trading rates are integrable in Subsection 6.7.3 but not in Subsection 6.7.4.

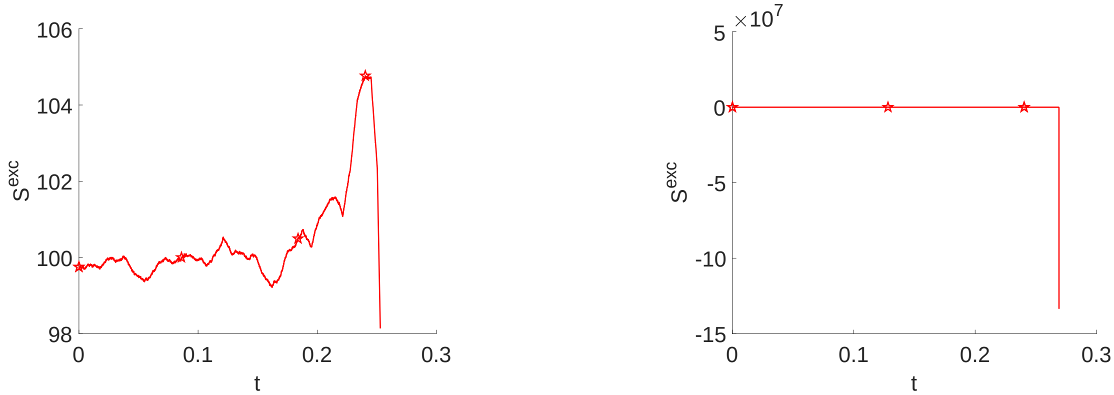

Figure 6.9: Depiction of the execution price for ω_{dn} in Subsection 6.7.4.

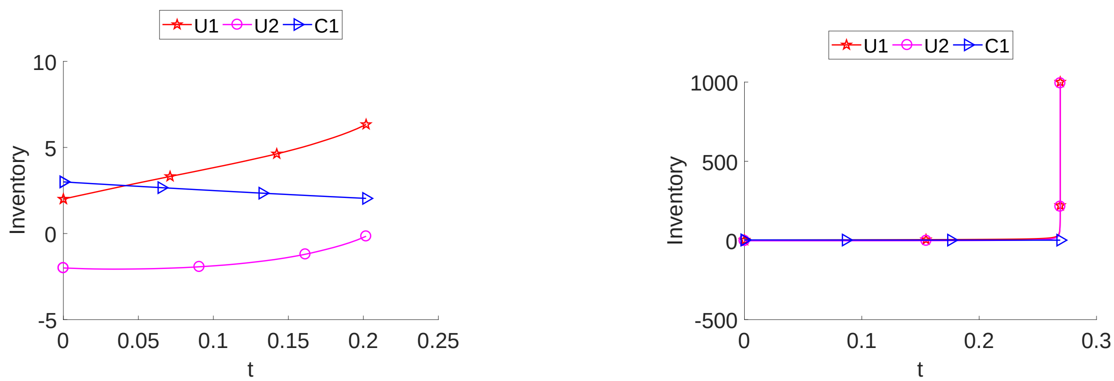

Figure 6.10: Depiction of the agents' inventories for ω_{up} in Subsection 6.7.4.

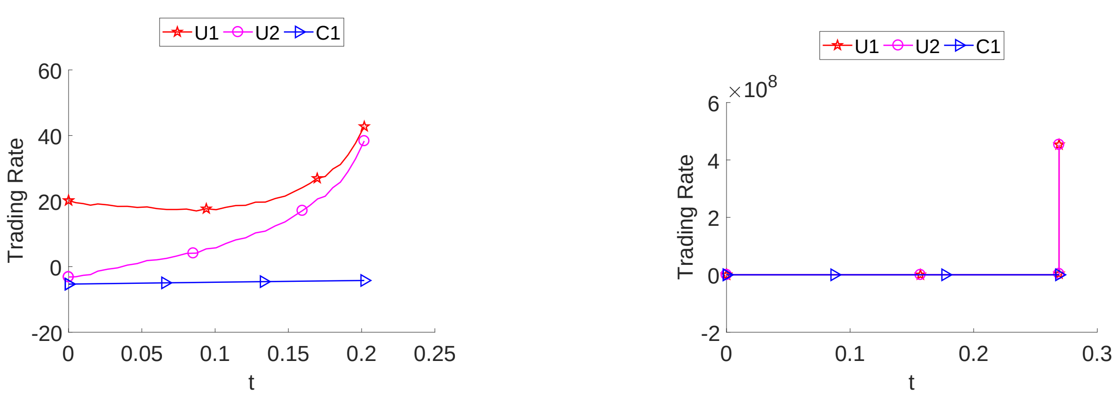

Figure 6.11: Depiction of the agents' trading rates for ω_{up} in Subsection 6.7.4.

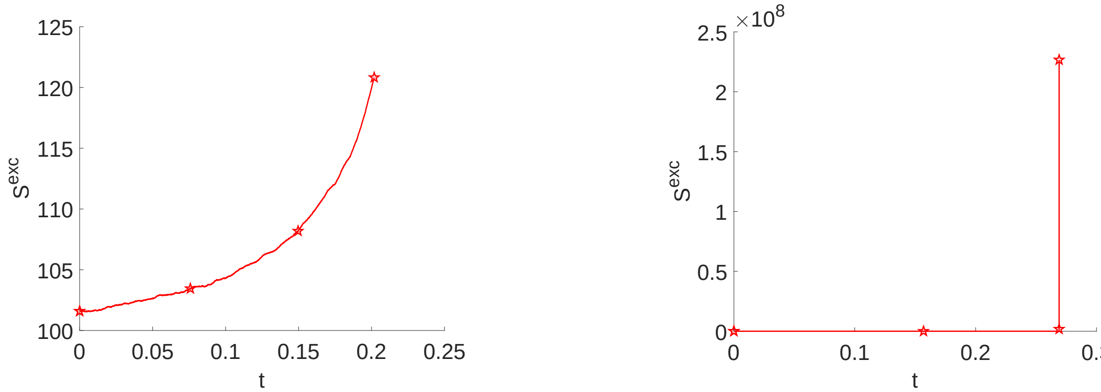

Figure 6.12: Depiction of the execution price for ω_{up} in Subsection 6.7.4.

6.8 Appendix: Section 6.4 Proofs

6.8.1 Proof of Lemma VI.9

We now implement the steps outlined in Remark VI.14.

Step 1: Denote the usual P_j-augmentation of $\left\{\mathcal{F}^{unf}_{j,t}\right\}_{0\leq t\leq T}$ by $\left\{\tilde{\mathcal{F}}^{unf}_{j,t}\right\}_{0\leq t\leq T}$. Let $\tilde{\mathcal{A}}_j$ be the space of $\tilde{\mathcal{F}}^{unf}_{j,t}$-progressively measurable processes θ_j such that (6.4) and (6.5) hold. We, again, define the process $X^{\theta_j}_j$ by (6.6) for any strategy $\theta_j \in \tilde{\mathcal{A}}_j$. Agent j's auxiliary problem is to maximize

$$E^{P_j}\left[-\int_0^T \theta_{j,t}\, S^{exc}_{j,\theta_j,t}\, dt - \frac{\kappa_j}{2}\int_0^T \left(X^{\theta_j}_{j,t}\right)^2 dt\right] \tag{6.33}$$

over $\theta_j \in \tilde{\mathcal{A}}_j$.

Step 2: We wish to show that

$$E^{P_j}\left[\beta_j \middle| \tilde{\mathcal{F}}^{unf}_{j,t}\right] = E^{P_j}\left[\beta_j \middle| \mathcal{F}^{unf}_{j,t}\right] \quad P_j - \text{a.s.}$$

$(\Omega_j, \mathcal{F}_j, P_j)$ is P_j-complete by hypothesis (see Subsection 6.4.2). Suppose that $t \in [0,T)$.[18] Letting $\mathcal{N}_j$ be the P_j-null subsets of Ω_j and

$$\mathcal{F}^{unf}_{j,t^+} = \bigcap_{t<u\leq T} \mathcal{F}^{unf}_{j,u},$$

[18]The result is clear when $t = T$, since $\tilde{\mathcal{F}}^{unf}_{j,T} = \sigma\left(\mathcal{F}^{unf}_{j,T}, \mathcal{N}_j\right)$.

we have

$$\tilde{\mathcal{F}}^{unf}_{j,t} = \sigma\left(\mathcal{F}^{unf}_{j,t^+}, \mathcal{N}_j\right).$$

Hence,

$$E^{P_j}\left[\beta_j \middle| \tilde{\mathcal{F}}^{unf}_{j,t}\right] = E^{P_j}\left[\beta_j \middle| \mathcal{F}^{unf}_{j,t^+}\right] \qquad P_j - \text{a.s.} \tag{6.34}$$

Since $\mathcal{F}^{unf}_{j,t} \subseteq \tilde{\mathcal{F}}^{unf}_{j,t}$, it suffices to show that

$$E^{P_j}\left[\mathbf{1}_U \beta_j\right] = E^{P_j}\left[\mathbf{1}_U E^{P_j}\left[\beta_j \left| \mathcal{F}^{unf}_{j,t}\right.\right]\right]$$

for all $U \in \mathcal{F}^{unf}_{j,t^+}$. Pick $U \in \mathcal{F}^{unf}_{j,t^+}$ and any positive decreasing sequence $(\epsilon_n)_{n\geq 1}$ in $(0, T-t)$ tending to 0. By (6.2) in Subsection 6.4.2 and (6.11) in Step 2,

$$\mathbf{1}_U E^{P_j}\left[\beta_j \left| \mathcal{F}^{unf}_{j,t+\epsilon_n}\right.\right] \xrightarrow{P_j-a.s.} \mathbf{1}_U E^{P_j}\left[\beta_j \left| \mathcal{F}^{unf}_{j,t}\right.\right]. \tag{6.35}$$

By the Vitali convergence theorem and the uniform integrability of the collection

$$\left\{\mathbf{1}_U E^{P_j}\left[\beta_j \left| \mathcal{F}^{unf}_{j,t+\epsilon_n}\right.\right]\right\}_{n\geq 1},$$

(6.35) also hold in the sense of L^1-convergence. This finishes the argument, as

$$E^{P_j}\left[\mathbf{1}_U \beta_j\right] = E^{P_j}\left[\mathbf{1}_U E^{P_j}\left[\beta_j \left| \mathcal{F}^{unf}_{j,t+\epsilon_n}\right.\right]\right]$$

for $n \geq 1$.

Step 3: By (6.3) and (6.6),

$$-\int_0^T \theta_{j,t}\, S^{exc}_{j,\theta_j,t}\, dt = -\int_0^T \theta_{j,t}\, S^{unf}_{j,t}\, dt - \int_0^T \theta_{j,t}\left[\eta_{j,per}\left(X^{\theta_j}_{j,t} - x_j\right) + \frac{1}{2}\eta_{j,tem}\theta_{j,t}\right] dt$$

for $\theta_j \in \tilde{\mathcal{A}}_j$. Section 7.4 of [162] and (6.2) imply that the process $\left\{\overline{W}_{j,t}\right\}_{0\leq t\leq T}$ with

$$\overline{W}_{j,t} \triangleq S^{unf}_{j,t} - S_{j,0} - \int_0^t E^{P_j}\left[\beta_j \middle| \mathcal{F}^{unf}_{j,s}\right] ds$$

is an $\mathcal{F}^{unf}_{j,t}$-Wiener process under P_j[19] and

$$S^{unf}_{j,t} = S_{j,0} + \int_0^t E^{P_j}\left[\beta_j \middle| \mathcal{F}^{unf}_{j,s}\right] ds + \overline{W}_{j,t}. \tag{6.36}$$

After integrating by parts and recalling (6.6) and (6.36), we get

$$\begin{aligned} &E^{P_j}\left[-\int_0^T \theta_{j,t}\, S^{unf}_{j,t}\, dt\right] \\ &\quad = E^{P_j}\left[-X^{\theta_j}_{j,T} S^{unf}_{j,T} + \int_0^T X^{\theta_j}_{j,t}\, E^{P_j}\left[\beta_j \middle| \mathcal{F}^{unf}_{j,t}\right] dt\right] + x_j S_{j,0}. \end{aligned}$$

We also have

$$\begin{aligned} &E^{P_j}\left[-\int_0^T \theta_{j,t}\left[\eta_{j,per}\left(X^{\theta_j}_{j,t} - x_j\right) + \frac{1}{2}\eta_{j,tem}\theta_{j,t}\right] dt\right] \\ &\quad = E^{P_j}\left[-\frac{1}{2}\eta_{j,per}\left(X^{\theta_j}_{j,T} - x_j\right)^2 - \frac{1}{2}\eta_{j,tem}\int_0^T \theta^2_{j,t}\, dt\right]. \end{aligned}$$

Now $X^{\theta_j}_{j,T} = 0$ P_j-a.s. by the definition of $\tilde{\mathcal{A}}_j$ in Step 1. Since x_j, $S_{j,0}$ and $S^{unf}_{j,T}$ do not depend on Agent j's choice of $\theta_j \in \tilde{\mathcal{A}}_j$, Step 2 implies that $\theta^\star_j$ maximizes (6.33) over $\theta_j \in \tilde{\mathcal{A}}_j$ if and only if it maximizes

$$E^{P_j}\left[\int_0^T X^{\theta_j}_{j,t}\, E^{P_j}\left[\beta_j \middle| \tilde{\mathcal{F}}^{unf}_{j,t}\right] dt - \frac{1}{2}\eta_{j,tem}\int_0^T \theta^2_{j,t}\, dt - \frac{\kappa_j}{2}\int_0^T \left(X^{\theta_j}_{j,t}\right)^2 dt\right]. \tag{6.37}$$

Due to (6.2), (6.11), in Step 2, and Agent j's Gaussian prior for β_j, the process $E^{P_j}\left[\beta_j \middle| \tilde{\mathcal{F}}^{unf}_{j,\cdot}\right]$ is $\tilde{\mathcal{F}}^{unf}_{j,t}$-predictable and in $L^2\left(dP_j \otimes dt\right)$. Clearly, $\theta^\star_j$ maximizes (6.37) over $\theta_j \in \tilde{\mathcal{A}}_j$ if and only if it minimizes

$$E^{P_j}\left[\frac{1}{2}\int_0^T \left(X^{\theta_j}_{j,t} - \frac{E^{P_j}\left[\beta_j \middle| \tilde{\mathcal{F}}^{unf}_{j,t}\right]}{\kappa_j}\right)^2 dt + \frac{\eta_{j,tem}}{2\kappa_j}\int_0^T \theta^2_{j,t}\, dt\right]. \tag{6.38}$$

[19] In fact, $\overline{W}_j$ is an *innovation process*, i.e., for each $t \in [0,T]$, we have $\mathcal{F}^{unf}_{j,t} = \mathcal{F}^{\overline{W}_j}_{j,t}$. Here, $\left\{\mathcal{F}^{\overline{W}_j}_{j,t}\right\}_{0 \le t \le T}$ is the filtration generated by $\overline{W}_j$.

Step 4: After defining

$$K_j(t,s) \triangleq \sqrt{\frac{\kappa_j}{\eta_{j,tem}}} \left(\frac{\sinh(\tau_j(s))}{\cosh(\tau_j(t)) - 1} \right), \qquad 0 \le t \le s < T$$

$$\hat{\beta}_{j,t} \triangleq E^{P_j}\left[\frac{1}{\kappa_j} \left(1 - \frac{1}{\cosh(\tau_j(t))} \right) \right. \tag{6.39}$$

$$\left. \cdot \int_t^T E^{P_j}\left[\beta_j \middle| \tilde{\mathcal{F}}^{unf}_{j,s} \right] K_j(t,s)\, ds \,\middle|\, \tilde{\mathcal{F}}^{unf}_{j,t} \right], \qquad t \in [0,T),$$

we see from Theorem 3.2 of [34] that (6.38) has a unique solution $\theta^\star_j \in \tilde{\mathcal{A}}_j$. Moreover, the corresponding optimal inventory process $X^{\theta^\star_j}_j$ satisfies the linear ODE

$$\begin{aligned} dX^{\theta^\star_j}_{j,t} &= \sqrt{\frac{\kappa_j}{\eta_{j,tem}}} \coth(\tau_j(t)) \left(\hat{\beta}_{j,t} - X^{\theta^\star_j}_{j,t} \right) dt \\ X^{\theta^\star_j}_{j,0} &= x_j \end{aligned} \tag{6.40}$$

$dP_j \otimes dt$-a.s. on $\Omega_j \times [0,T)$.

Using Fubini's theorem and Steps 2, we get that

$$E^{P_j}\left[\int_t^T E^{P_j}\left[\beta_j \middle| \tilde{\mathcal{F}}^{unf}_{j,s} \right] K_j(t,s)\, ds \,\middle|\, \tilde{\mathcal{F}}^{unf}_{j,t} \right] = E^{P_j}\left[\beta_j \middle| \mathcal{F}^{unf}_{j,t} \right], \quad P_j - \text{a.s.} \tag{6.41}$$

The tanh half-angle formula together with (6.11) and (6.39) imply that (6.40) can be re-written as

$$\begin{aligned} \theta^\star_{j,t} &= -\sqrt{\frac{\kappa_j}{\eta_{j,tem}}} \coth(\tau_j(t))\, X^{\theta^\star_j}_{j,t} \\ &\quad + \frac{\tanh(\tau_j(t)/2) \left[\mu_j + \nu_j^2 \left(S^{unf}_{j,t} - S_{j,0} \right) \right]}{\sqrt{\eta_{j,tem}\kappa_j} \left(1 + \nu_j^2 t \right)}, \quad t \in (0,T) \\ X^{\theta^\star_j}_{j,0} &= x_j. \end{aligned} \tag{6.42}$$

Step 5: We know that $\theta^\star_j$ satisfies (6.4) and (6.5), as all strategies in $\tilde{\mathcal{A}}_j$ have these properties. Now $W_{j,\cdot}(\omega)$ is continuous on $[0,T]$ for P_j-almost every $\omega \in \Omega_j$.

When such an ω is chosen, (6.42) becomes (6.10). The latter is a first order linear ODE with continuous coefficients, so $\theta^{\star}_{j,\cdot}(\omega)$ is continuous on $[0,T)$ (e.g., by Chapter 1.2 of [231]).

Since our terminal inventory constraint is deterministic, we observe that

$$\lim_{t\uparrow T}\theta^{\star}_{j,t}(\omega)$$

exists and is finite from (28) and (29) in the proof of Theorem 3.2 in [34], as well as (6.41) in Step 4. In particular, we can view the paths of $\theta^{\star}_{j}$ on $[0,T]$ as P_j-a.s. continuous.[20] We conclude by noting that $\theta^{\star}_{j}$ is also $\mathcal{F}^{unf}_{j,t}$-adapted by (28) and (29) in the proof of Theorem 3.2 in [34], (6.11) in Step 2, and (6.41) in Step 4.

$\square$

6.9 Appendix: Section 6.6 Proofs

6.9.1 Proof of Lemma VI.23

Let $j\in\{1,\ldots,K\}$. At each time t, Agent j observes the correct value of $S^{exc}_{t}(\omega)$, interprets this value as the realized value of $S^{exc}_{j,\theta^{\star}_{j},t}(\omega)$, and computes $S^{unf}_{j,t}(\omega)$.[21] By (6.3), it follows that

$$\begin{aligned} S^{exc}_{t}(\omega) &= S^{exc}_{j,\theta^{\star}_{j},t}(\omega) \\ &= S^{unf}_{j,t}(\omega)+\eta_{j,per}\left(X^{\theta^{\star}_{j}}_{j,t}(\omega)-x_j\right)+\frac{1}{2}\eta_{j,tem}\theta^{\star}_{j,t}(\omega). \end{aligned} \tag{6.43}$$

[20]Alternatively, we could give an argument using singular point theory as in Section 6.6.

[21]By abuse of notation, we evaluate $S^{exc}_{j,\theta^{\star}_{j},t}$ and $S^{unf}_{j,t}$ are evaluated at ω; however, Agent j would evaluate these quantities at some $\omega_j\in\Omega_j$. We adopt similar conventions in the sequel without further comment.

After substituting (6.13) into (6.43), we have

$$\begin{aligned}
&S_{j,t}^{unf}(\omega) - S_{j,0} \\
&\quad = (S_0 - S_{j,0}) + \tilde{\beta} t + \sum_{\substack{i \leq K \\ i \neq j}} \tilde{\eta}_{i,per}\left(X_{i,t}^{\theta_i^\star}(\omega) - x_i\right) + \sum_{i>K} \tilde{\eta}_{i,per}\left(X_{i,t}^{\theta_i^\star} - x_i\right) \\
&\qquad + \frac{1}{2}\sum_{\substack{i \leq K \\ i \neq j}} \tilde{\eta}_{i,tem}\theta_{i,t}^\star(\omega) + \frac{1}{2}\sum_{i>K} \tilde{\eta}_{i,tem}\theta_{i,t}^\star \\
&\qquad + (\tilde{\eta}_{j,per} - \eta_{j,per})\left(X_{j,t}^{\theta_j^\star}(\omega) - x_j\right) + \frac{1}{2}(\tilde{\eta}_{j,tem} - \eta_{j,tem})\,\theta_{j,t}^\star(\omega) + \tilde{W}_t(\omega). \quad (6.44)
\end{aligned}$$

The quantity on the LHS of (6.44) plays a role in determining Agent j's strategy (see Lemma VI.9). Substituting (6.44) into (6.10) and applying the half-angle formula for $\tanh(\cdot)$, we get

$$\begin{aligned}
&A_{jj}(t)\,\theta_{j,t}^\star(\omega) - \sum_{\substack{i \leq K \\ i \neq j}} A_{ji}(t)\,\theta_{i,t}^\star(\omega) \\
&\quad = B_{jj}(t)\,X_{j,t}^{\theta_j^\star}(\omega) + \sum_{\substack{i \leq K \\ i \neq j}} B_{ji}(t)\,X_{i,t}^{\theta_i^\star}(\omega) + C_j(t,\omega).
\end{aligned}$$

It follows that the uncertain agents' strategies are characterized by the ODE system

$$\begin{aligned}
A(t)\,\theta_t^{u,\star}(\omega) &= B(t)\,X_t^{u,\theta^\star}(\omega) + C(t,\omega) \\
X_0^{u,\theta^\star}(\omega) &= x^u. \qquad (6.45)
\end{aligned}$$

Corollary VI.15, Lemma VI.21 and a standard existence and uniqueness theorem (see Sections 1.1 and 3.1 of [36]) finish the argument. □

6.9.2 Proof of Lemma VI.24

As $t \uparrow t_e$,

$$\left\{A(t)\,\dot{X}_t^u(\omega) = B(t)\,X_t^u(\omega)\right\}$$

$$\Longleftrightarrow$$

$$\left\{[\det A(t)]\,\dot{X}_t^u(\omega) = [\mathrm{adj} A(t)]\,B(t)\,X_t^u(\omega)\right\}.$$

Here, adj denotes the usual adjugate operator.

We can find a non-negative integer m such that the multiplicity of the zero of $\det A$ at t_e is $(m+1)$ by Lemma VI.21. Hence, there is a unique non-vanishing analytic function f such that

$$\det A(t) = (t - t_e)^{m+1} f(t) \tag{6.46}$$

on a small neighborhood of t_e. Note that f is non-vanishing, as the zeroes of $\det A$ are isolated and $\det A(T) = 1$ (see Lemma VI.21). We then define the analytic (see Lemma VI.21) map D by

$$D(t) \triangleq [\mathrm{adj} A(t)] B(t) / f(t) \tag{6.47}$$

and arrive at (6.16).

Since $\det A(\cdot)$ has a root at t_e, the rank of $A(t_e)$ is no more than $K-1$. We conclude by observing that adj $A(t_e)$ has rank 1 when $A(t_e)$ has rank $K-1$; otherwise, adj $A(t_e)$ must be the zero matrix. The comments about the rank of $D(t_e)$ immediately follow. □

6.9.3 Proof of Lemma VI.27

$D(t_e) \neq 0$ since $\lambda \neq 0$. Then (6.15) has a singular point of the first kind at t_e (see our discussion above). $\lambda \notin \mathbb{Z}$ by hypothesis, so Theorem 6.5 of [87] implies that a fundamental solution of (6.15) on $[t_e - \rho, t_e)$ for some $\rho > 0$ is given by

$$P(t) |t - t_e|^{D(t_e)}. \tag{6.48}$$

In (6.48), $P(\cdot)$ is an analytic $M_K(\mathbb{R})$-valued function with $P(t_e) = I_K$. Moreover, $P(t)$ is invertible for all $t \in [t_e - \rho, t_e)$ and[22]

$$\left(P(t) |t - t_e|^R\right)^{-1} = |t - t_e|^{-R} [P(t)]^{-1}. \tag{6.49}$$

[22] Any fundamental solution of (6.16) is invertible everywhere, as are matrix exponentials.

The solution of (6.14) satisfies

$$(t-t_e)\,\theta_t^{u,\star}(\omega) = D(t)\,X_t^{u,\theta^\star}(\omega) + \frac{\operatorname{adj}[A(s)]\,C(s,\omega)}{f(s)}.$$

near t_e (argue as in Lemma VI.24). Since

$$P(t)\,|t-t_e|^{D(t_e)}\,\rho^{-D(t_e)}\,[P(t_e-\rho)]^{-1}$$

is also a fundamental solution of (6.16) on $[t_e-\rho, t_e)$[23] and equals I_K at $t_e-\rho$, we can apply variation of parameters[24] to obtain

$$\begin{aligned} &X_t^{u,\theta^\star}(\omega) \\ &= P(t)\,|t-t_e|^{D(t_e)}\left[\rho^{-D(t_e)}\,[P(t_e-\rho)]^{-1}\right]\cdot\Bigg[X_{t_e-\rho}^{\theta^\star}(\omega) \\ &\quad + \int_{t_e-\rho}^{t}\left(P(t_e-\rho)\,\rho^{D(t_e)}\,|s-t_e|^{-D(t_e)}\,[P(s)]^{-1}\right)\left(\frac{\operatorname{adj}[A(s)]\,C(s,\omega)}{(s-t_e)\,f(s)}\right)ds\Bigg]. \end{aligned} \tag{6.50}$$

We can find an eigenbasis $\{v_1,\ldots,v_K\}$ for $D(t_e)$ such that v_K corresponds to λ (see Lemma VI.24 and Remark VI.26). We then define the continuous real-valued functions $\{F_1(\cdot,\omega),\ldots,F_K(\cdot,\omega)\}$ on $[t_e-\rho,t_e]$ and the constants $\{y_1(\omega),\ldots,y_K(\omega)\}$ as certain eigenbasis coordinates:

$$\begin{aligned} \sum_{j=1}^{K} F_j(s,\omega)\,v_j &\triangleq \frac{[P(s)]^{-1}\operatorname{adj}[A(s)]\,C(s,\omega)}{f(s)} \\ \sum_{j=1}^{K} y_j(\omega)\,v_j &\triangleq \rho^{-D(t_e)}\,[P(t_e-\rho)]^{-1}\,X_{t_e-\rho}^{\theta^\star}(\omega). \end{aligned} \tag{6.51}$$

Taken with (6.50), these definitions immediately give (6.17) after recalling that for any matrix $Q\in M_K(\mathbb{R})$ with eigenvalue γ and corresponding eigenvector v, we have

$$|t-t_e|^{Q}\,v = |t-t_e|^{\gamma}\,v.$$

$\square$

[23]See Theorem 2.5 of Coddington & Carlson ([87]).
[24]See Theorem 2.8 of Coddington & Carlson ([87]).

6.9.4 Proof of Lemma VI.28

We know that $S^{exc}(\omega)$, the $X_j^{\theta_j^\star}(\omega)$'s and the $\theta_j^\star(\omega)$'s are all uniquely defined and continuous on $[0,T)$ (see Lemma VI.23). Corollary VI.15 implies that $X_j^{\theta_j^\star}(\omega)$ and $\theta_j^\star(\omega)$ are continuous at T for $j>K$ (the certain agents). It also gives us

$$\lim_{t\uparrow T} X_{j,t}^{\theta_j^\star}(\omega)=0$$

for $j>K$. By Definition VI.17, it remains to show that

$$\lim_{t\uparrow T} X_t^{u,\theta^\star}(\omega)=0 \quad \text{and} \quad \lim_{t\uparrow T}\theta_t^{u,\star}(\omega)\in\mathbb{R}^K. \tag{6.52}$$

As discussed above, one difficulty is that the diagonal entries of B in (6.14) explode at T (see Lemma VI.21); however, the approach for resolving this issue is similar to that used to analyze solution behavior near t_e.

First, we show that (6.15) (after replacing t_e with T) has a singular point of the first kind at T. Now $\sinh(\tau_j(\cdot))$ has a zero of multiplicity 1 at T since

$$\left.\frac{d\sinh(\tau_j(t))}{dt}\right|_{t=T}=-\sqrt{\frac{\kappa_j}{\eta_{j,tem}}}\cosh(\tau_j(t))\Bigg|_{t=T}=-\sqrt{\frac{\kappa_j}{\eta_{j,tem}}}.$$

Hence, there is a unique non-vanishing analytic function g_j such that

$$\sinh(\tau_j(t))=(t-T)\,g_j(t) \quad \text{and} \quad g_j(T)=-\sqrt{\frac{\kappa_j}{\eta_{j,tem}}} \tag{6.53}$$

on a small neighborhood of T. Near T, it follows that the entries of $(t-T)\,B(t)$ are given by

$$(t-T)\,B_{ik}(t)=\begin{cases}(t-T)\left(\tilde{\eta}_{i,per}-\eta_{i,per}\right)\Phi_i(t) \\ \qquad -\sqrt{\dfrac{\kappa_i}{\eta_{i,tem}}}\left(\dfrac{\cosh(\tau_i(t))}{g_i(t)}\right) & \text{if } i=k \\ \\ (t-T)\,\tilde{\eta}_{k,per}\Phi_i(t) & \text{if } i\neq k\end{cases} \tag{6.54}$$

(see Definition VI.19). On this region, the solution of (6.15) satisfies

$$(t-T)\,\dot{X}_t^u(\omega) = A^{-1}(t)\,(t-T)\,B(t)\,X_t^u(\omega)\,. \tag{6.55}$$

By (6.54) and Lemma VI.21, (6.55) has a singular point of the first kind at T.

Second, we find a fundamental solution of (6.55) near T. We know that

$$A^{-1}(T) = (t-T)\,B(t)\Big|_{t=T} = I_K$$

by (6.53), (6.54), and Lemma VI.21. Theorem 6.5 of [87] implies that a fundamental solution of (6.55) on $[T-\delta, T)$ for some $\delta > 0$ is given by

$$Q(t)\,|t-T|^{I_K} = Q(t)\,|t-T|\,. \tag{6.56}$$

In (6.56), Q is an analytic $M_K(\mathbb{R})$-valued function with $Q(T) = I_K$. Also, $Q(t)$ is invertible for all $t \in [T-\delta, T)$.[25]

Finally, we use our fundamental solution to solve (6.14) and conclude the proof. Notice that $\tanh(\tau_j(\cdot))$ also has a zero of multiplicity 1 at T since

$$\frac{d\tanh(\tau_j(t))}{dt}\Bigg|_{t=T} = -\frac{1}{2}\sqrt{\frac{\kappa_j}{\eta_{j,tem}}}\operatorname{sech}^2(\tau_j(t)/2)\Bigg|_{t=T} = -\frac{1}{2}\sqrt{\frac{\kappa_j}{\eta_{j,tem}}}.$$

There is a unique non-vanishing analytic function h_j such that

$$\tanh(\tau_j(t)/2) = (t-T)\,h_j(t) \tag{6.57}$$

on a neighborhood of T. In particular, the entries of $C(t,\omega)/(t-T)$ near T are given by

$$\begin{aligned}\frac{C_i(t,\omega)}{(t-T)} = \left(\frac{h_i(t)\,\nu_i^2}{\sqrt{\eta_{i,tem}\kappa_i}\,(1+\nu_i^2 t)}\right)\Bigg[&\frac{\mu_i}{\nu_i^2} + (S_0 - S_{i,0}) + \tilde{\beta}t - \sum_{\substack{k\le K\\ k\ne i}}\tilde{\eta}_{k,per}x_k \\ &- x_i(\tilde{\eta}_{i,per} - \eta_{i,per}) + \sum_{k>K}\tilde{\eta}_{k,per}\left(X_{k,t}^{\theta_k^\star} - x_k\right) \\ &+\frac{1}{2}\sum_{k>K}\tilde{\eta}_{k,tem}\theta_{k,t}^\star + \tilde{W}_t(\omega)\Bigg].\end{aligned} \tag{6.58}$$

[25] Any fundamental solution of (6.55) is invertible everywhere.

Since

$$Q\left(t\right)\left|t-T\right|\delta^{-1}Q^{-1}\left(T-\delta\right)$$

is also a fundamental solution of (6.16) on $[T-\delta, T)$[26] and equals I_K at $T-\delta$, we can apply variation of parameters[27] to obtain

$$\begin{aligned} & X_t^{u,\theta^\star}\left(\omega\right) \\ & \quad = Q\left(t\right)\left|t-T\right|\delta^{-1}Q^{-1}\left(T-\delta\right)\cdot\left[X_{T-\delta}^{\theta^\star}\left(\omega\right)\right. \\ & \quad\quad \left. + \int_{T-\delta}^{t}\left(Q\left(T-\delta\right)\delta\left|s-T\right|^{-1}Q^{-1}\left(s\right)\right)A^{-1}\left(s\right)C\left(s,\omega\right)ds\right]. \end{aligned} \tag{6.59}$$

By (6.58), (6.59), and Corollary VI.15, we get (6.52). □

6.10 Appendix: Section 6.7 Proofs

6.10.1 Proof of Lemma VI.34

By Definitions VI.19 and VI.32, we see that A is now given by

$$A_{ik}\left(t\right) \triangleq \begin{cases} 1-\frac{1}{2}\left(\tilde{\eta}_{tem}-\eta_{tem}\right)\Phi\left(t\right) & \text{if } i=k \\ -\frac{1}{2}\tilde{\eta}_{tem}\Phi\left(t\right) & \text{if } i\neq k \end{cases}. \tag{6.60}$$

A short calculation shows that

$$\det A\left(t\right) = \left[1+\frac{1}{2}\eta_{tem}\Phi\left(t\right)\right]^{K-1}\left[1-\frac{1}{2}\left(K\tilde{\eta}_{tem}-\eta_{tem}\right)\Phi\left(t\right)\right]. \tag{6.61}$$

The first term in (6.61) is always at least 1. The second term is non-zero at 0 but does have a root on $(0, T]$ if and only if (6.19) holds.[28] Both of these observations come from Lemma VI.21.

Now, (6.19) implies that $K\tilde{\eta}_{tem} > \eta_{tem}$. Since t_e is a zero of $\det A$, we must have that

$$1-\frac{1}{2}\left(K\tilde{\eta}_{tem}-\eta_{tem}\right)\Phi\left(t_e\right)=0. \tag{6.62}$$

[26] See Theorem 2.5 of Coddington & Carlson ([87]).
[27] See Theorem 2.8 of Coddington & Carlson ([87]).
[28] In fact, t_e is the unique root of $\det A$ in this case.

Hence, by Lemma VI.21,

$$\left.\frac{d\left[\det A\left(t\right)\right]}{dt}\right|_{t=t_e} = -\frac{1}{2}\left(K\tilde{\eta}_{tem}-\eta_{tem}\right)\left[1+\frac{1}{2}\eta_{tem}\Phi\left(t\right)\right]^{K-1}\left.\dot{\Phi}\left(t\right)\right|_{t=t_e} > 0. \quad (6.63)$$

□

6.10.2 Proof of Lemma VI.37

By (6.46), (6.63), and Lemma VI.34,

$$\begin{aligned} f\left(t_e\right) &= \left.\frac{d\left[\det A\left(t\right)\right]}{dt}\right|_{t=t_e} \\ &= -\frac{1}{2}\left(K\tilde{\eta}_{tem}-\eta_{tem}\right)\left[1+\frac{1}{2}\eta_{tem}\Phi\left(t_e\right)\right]^{K-1}\dot{\Phi}\left(t_e\right). \end{aligned} \quad (6.64)$$

A short calculation shows that adj $A\left(t\right)$ is given by

$$\begin{aligned} &\left[\mathrm{adj}A\left(t\right)\right]_{ik} \\ &= \left(1+\frac{1}{2}\eta_{tem}\Phi\left(t\right)\right)^{K-2}\begin{cases} 1-\frac{1}{2}\left[\left(K-1\right)\tilde{\eta}_{tem}-\eta_{tem}\right]\Phi\left(t\right) & \text{if } i=k \\ \frac{1}{2}\tilde{\eta}_{tem}\Phi\left(t\right) & \text{if } i\neq k \end{cases}. \end{aligned} \quad (6.65)$$

It follows that

$$\begin{aligned} &\left[\left(\mathrm{adj}A\left(t\right)\right]B\left(t\right)\right]_{ik} \\ &= \tilde{\eta}_{per}\Phi\left(t\right)\left(1+\frac{1}{2}\eta_{tem}\Phi\left(t\right)\right)^{K-1} \\ &\quad+\left(1+\frac{1}{2}\eta_{tem}\Phi\left(t\right)\right)^{K-2}\left(\eta_{per}\Phi\left(t\right)+\sqrt{\frac{\kappa}{\eta_{tem}}}\coth\left(\tau\left(t\right)\right)\right) \\ &\quad\cdot\begin{cases} \frac{1}{2}\left[\left(K-1\right)\tilde{\eta}_{tem}-\eta_{tem}\right]\Phi\left(t\right)-1 & \text{if } i=k \\ \\ -\frac{1}{2}\tilde{\eta}_{tem}\Phi\left(t\right) & \text{if } i\neq k \end{cases}. \end{aligned} \quad (6.66)$$

One can then check that the only potentially non-zero eigenvalue of

$$D\left(t_e\right) = \frac{\left[\mathrm{adj}A\left(t_e\right)\right]B\left(t_e\right)}{f\left(t_e\right)}$$

is given by

$$\lambda = -\frac{2\left[\left(K\tilde{\eta}_{per} - \eta_{per}\right)\Phi\left(t_e\right) - \sqrt{\frac{\kappa}{\eta_{tem}}}\coth\left(\tau\left(t_e\right)\right)\right]}{\left(K\tilde{\eta}_{tem} - \eta_{tem}\right)\dot{\Phi}\left(t_e\right)} \tag{6.67}$$

with corresponding eigenvector v_K as above. We get (6.22) from (6.67) after applying (6.21).

Recall that $\Phi\left(t_e\right) > 0$ and $\dot{\Phi}\left(t_e\right) < 0$ by Lemma VI.21. Since t_e, Φ, and τ do not depend on $\tilde{\eta}_{per}$ or η_{per}, we can ensure that $\lambda \notin \mathbb{Z}$ by perturbing the latter parameters. $D\left(t_e\right)$ is then diagonalizable as observed in Lemma VI.27, and $v_1, \ldots, v_{K-1}$ can be computed using (6.66). □

6.10.3 Proof of Lemmas VI.39 and VI.42

Since our uncertain agents are semi-symmetric,

$$\begin{aligned}
C_i\left(t,\omega\right) = \Phi\left(t\right)\tilde{W}_t\left(\omega\right) \\
&+ \Phi\left(t\right)\left[\tilde{\beta}t + \sum_{k>K}\tilde{\eta}_{k,per}\left(X_{k,t}^{\theta_k^\star} - x_k\right) + \frac{1}{2}\sum_{k>K}\tilde{\eta}_{k,tem}\theta_{k,t}^\star\right] \\
&+ \Phi\left(t\right)\left[\frac{\mu_i}{\nu^2} + \left(S_0 - S_{i,0}\right) - \sum_{\substack{k\leq K\\ k\neq i}}\tilde{\eta}_{per}x_k - x_i\left(\tilde{\eta}_{per} - \eta_{per}\right)\right]
\end{aligned} \tag{6.68}$$

for $t \leq t_e$ by Definition VI.19. For convenience, we introduce the following deterministic function[29] c and the constants $c_1, \ldots, c_K$:

$$c\left(t\right) \triangleq \left[\tilde{\beta}t + \sum_{k>K}\tilde{\eta}_{k,per}\left(X_{k,t}^{\theta_k^\star} - x_k\right) + \frac{1}{2}\sum_{k>K}\tilde{\eta}_{k,tem}\theta_{k,t}^\star\right]$$

$$\sum_{i=1}^{K} c_i v_i \triangleq \begin{bmatrix} \frac{\mu_1}{\nu^2} + \left(S_0 - S_{1,0}\right) - \sum_{\substack{k\leq K\\ k\neq 1}}\tilde{\eta}_{per}x_k - x_1\left(\tilde{\eta}_{per} - \eta_{per}\right) \\ \vdots \\ \frac{\mu_K}{\nu^2} + \left(S_0 - S_{K,0}\right) - \sum_{\substack{k\leq K\\ k\neq K}}\tilde{\eta}_{per}x_k - x_K\left(\tilde{\eta}_{per} - \eta_{per}\right) \end{bmatrix}. \tag{6.69}$$

[29]The function c is deterministic by Corollary VI.15.

Using (6.68), we get that

$$C(t,\omega) = \tilde{W}_t(\omega)\,\Phi(t)\,v_K + c(t)\,\Phi(t)\,v_K + \Phi(t)\sum_{i=1}^{K} c_i v_i. \tag{6.70}$$

By (6.65), $\{v_1, \ldots, v_K\}$ is an eigenbasis for $\mathrm{adj}\,[A(t)]$. Moreover,

$$\left(1 + \frac{1}{2}\eta_{tem}\Phi(t)\right)^{K-2}\left[1 - \frac{1}{2}\left(K\tilde{\eta}_{tem} - \eta_{tem}\right)\Phi(t)\right] \tag{6.71}$$

is the eigenvalue corresponding to each of $v_1, \ldots, v_{K-1}$, while

$$\left(1 + \frac{1}{2}\eta_{tem}\Phi(t)\right)^{K-1} \tag{6.72}$$

corresponds to v_K.

By (6.51), it follows that

$$\begin{aligned}
&\sum_{j=1}^{K} F_j(t,\omega)\,v_j \\
&\quad= \frac{[P(t)]^{-1}\,\mathrm{adj}\,[A(t)]\,C(t,\omega)}{f(t)} \\
&\quad= \tilde{W}_t(\omega)\left(\frac{\Phi(t)\left(1 + \frac{1}{2}\eta_{tem}\Phi(t)\right)^{K-1}}{f(t)}\right)[P(t)]^{-1}\,v_K \\
&\qquad\quad+ \left(\frac{\Phi(t)\left(1 + \frac{1}{2}\eta_{tem}\Phi(t)\right)^{K-1}\left(c(t) + c_K\right)}{f(t)}\right)[P(t)]^{-1}\,v_K \\
&\qquad\quad+ \left(\frac{\Phi(t)\left(1 + \frac{1}{2}\eta_{tem}\Phi(t)\right)^{K-2}\left[1 - \frac{1}{2}\left(K\tilde{\eta}_{tem} - \eta_{tem}\right)\Phi(t)\right]}{f(t)}\right) \\
&\qquad\qquad\cdot \sum_{i=1}^{K-1} c_i\,[P(t)]^{-1}\,v_i.
\end{aligned} \tag{6.73}$$

It follows that we can find analytic deterministic functions $F_{j,1}$ and $F_{j,2}$ such that

$$F_j(t,\omega) \triangleq \tilde{W}_t(\omega)\,F_{j,1}(t) + F_{j,2}(t) \tag{6.74}$$

for each $j \in \{1, \ldots, K\}$.[30] Since $P(t_e) = I_K$ (see Lemma VI.27), (6.64) and Remark

[30] Note that c is continuously differentiable on $[0, t_e]$ by Corollary VI.15.

VI.36 further imply that

$$F_{j,1}\left(t_e\right) = F_{j,2}\left(t_e\right) = \cdots = F_{K-1,1}\left(t_e\right) = F_{K-1,2}\left(t_e\right) = 0 \tag{6.75}$$

and

$$F_{K,1}\left(t_e\right) = -\frac{\Phi^2\left(t_e\right)}{\dot{\Phi}\left(t_e\right)} > 0 \quad \text{and} \quad F_{K,2}\left(t_e\right) = -\frac{\Phi^2\left(t_e\right)}{\dot{\Phi}\left(t_e\right)}\left(c\left(t_e\right) + c_K\right). \tag{6.76}$$

While $F_{K,1}\left(t_e\right) > 0$, determining the sign of $F_{K,2}\left(t_e\right)$ is difficult, in general, as it depends upon the sign of $c\left(t_e\right) + c_K$ (see (6.69)).

We see from (6.74) and (6.75) that the expression

$$\frac{F_j\left(s,\omega\right)}{\left|s - t_e\right|} \tag{6.77}$$

is bounded near t_e for each $j < K$ and almost every $\omega \in \tilde{\Omega}$. In particular, the coordinates of both

$$\sum_{j=1}^{K-1}\left(y_j\left(\omega\right) - \int_{t_e-\rho}^{t}\frac{F_j\left(s,\omega\right)}{\left|s - t_e\right|}\,ds\right)P\left(t\right)v_j$$

and its time derivative are bounded near t_e for such ω as well.

Since $P\left(t_e\right) = I_K$, the v_K-coordinate of $P\left(t\right)v_K$ tends to 1 $t \uparrow t_e$. For $j < K$, the v_j-coordinate of $P\left(t\right)v_K$ tends to 0 as $t \uparrow t_e$. In each situation, we can also obtain Lipschitz bounds on the convergence. Due to (6.17) and (6.74), potential explosions in the coordinates of $X_t^{u,\theta^\star}\left(\omega\right)$ are characterized by

$$\lim_{t\uparrow t_e}\left[\left|t - t_e\right|^{\lambda}\left(y_K\left(\omega\right) - \int_{t_e-\rho}^{t}\frac{\tilde{W}_s\left(\omega\right)F_{K,1}\left(s\right) + F_{K,2}\left(s\right)}{\left|s - t_e\right|^{1+\lambda}}\,ds\right)\right]. \tag{6.78}$$

Specifically,

$$\begin{aligned}
\{\left|(6.78)\right| < +\infty\} \quad &\Longleftrightarrow \quad \left\{\lim_{t\uparrow t_e} X_t^{u,\theta^\star}\left(\omega\right) \quad \text{exists in } \mathbb{R}^K\right\}\\
\{(6.78) = +\infty\} \quad &\Longleftrightarrow \quad \left\{\lim_{t\uparrow t_e} X_t^{u,\theta^\star}\left(\omega\right) = \left[+\infty,\ldots,+\infty\right]^\top\right\}\\
\{(6.78) = -\infty\} \quad &\Longleftrightarrow \quad \left\{\lim_{t\uparrow t_e} X_t^{u,\theta^\star}\left(\omega\right) = \left[-\infty,\ldots,-\infty\right]^\top\right\}.
\end{aligned} \tag{6.79}$$

To finish the proof, we separately consider the $\lambda < 0$ and $\lambda > 0$ cases.

$\lambda < 0$ **Case.**

Assume that $\lambda < 0$. It follows that

$$\lim_{t \uparrow t_e} \int_{t_e - \rho}^{t} \frac{\left|\tilde{W}_s(\omega) F_{K,1}(s)\right|}{\left|s - t_e\right|^{1+\lambda}} ds < \infty \quad \text{and} \quad \lim_{t \uparrow t_e} \int_{t_e - \rho}^{t} \frac{\left|F_{K,2}(s)\right|}{\left|s - t_e\right|^{1+\lambda}} ds < \infty.$$

Clearly,

$$\lim_{t \uparrow t_e} \left|t - t_e\right|^{\lambda} = +\infty,$$

meaning that

$$\begin{aligned} &\left\{ y_K(\omega) - \lim_{t \uparrow t_e} \int_{t_e - \rho}^{t} \frac{F_{K,2}(s)}{\left|s - t_e\right|^{1+\lambda}} ds > \lim_{t \uparrow t_e} \int_{t_e - \rho}^{t} \frac{\tilde{W}_s(\omega) F_{K,1}(s)}{\left|s - t_e\right|^{1+\lambda}} ds \right\} \\ &\quad \Longrightarrow \quad \left\{ \lim_{t \uparrow t_e} X_t^{u,\theta^\star}(\omega) = \left[+\infty, \ldots, +\infty\right]^\top \right\} \end{aligned} \tag{6.80}$$

and

$$\begin{aligned} &\left\{ y_K(\omega) - \lim_{t \uparrow t_e} \int_{t_e - \rho}^{t} \frac{F_{K,2}(s)}{\left|s - t_e\right|^{1+\lambda}} ds < \lim_{t \uparrow t_e} \int_{t_e - \rho}^{t} \frac{\tilde{W}_s(\omega) F_{K,1}(s)}{\left|s - t_e\right|^{1+\lambda}} ds \right\} \\ &\quad \Longrightarrow \quad \left\{ \lim_{t \uparrow t_e} X_t^{u,\theta^\star}(\omega) = \left[-\infty, \ldots, -\infty\right]^\top \right\} \end{aligned} \tag{6.81}$$

Arguing as in our discussion of (6.77), we see that the hypotheses in (6.80) and (6.81) also imply that

$$\left\{ \lim_{t \uparrow t_e} \theta_t^{u,\star}(\omega) = \left[+\infty, \ldots, +\infty\right]^\top \right\} \quad \text{and} \quad \left\{ \lim_{t \uparrow t_e} \theta_t^{u,\star}(\omega) = \left[-\infty, \ldots, -\infty\right]^\top \right\},$$

respectively.[31] Conditional on $\tilde{\mathcal{F}}_{t_e - \rho}$, the RHS of the inequality in (6.80) (and 6.81) is deterministic. Since $F_{K,1}(t_e) > 0$ (see (6.76)), we finish our proof of Lemma VI.39.

$\lambda > 0$ **Case.**

Assume that $\lambda > 0$. We can find a constant $R_0(\omega)$ such that

$$\left| y_K(\omega) - \int_{t_e - \rho}^{t} \frac{\tilde{W}_s(\omega) F_{K,1}(s) + F_{K,2}(s)}{\left|s - t_e\right|^{1+\lambda}} ds \right| \leq \frac{R_0(\omega)}{\left|t - t_e\right|^{\lambda}}. \tag{6.82}$$

[31] In particular, the coordinates of $\theta_t^{u,\star}(\omega)$ will asymptotically explode at the rate $\left|t - t_e\right|^{-\lambda-1}$.

Hence, (6.78) is bounded as $t \uparrow t_e$ and

$$\lim_{t \uparrow t_e} X_t^{u,\theta^\star}(\omega)$$

exists in $\mathbb{R}^K$ by our previous comments.

By our discussion surrounding (6.77), we see that explosions in the coordinates of $\theta_t^{u,\star}(\omega)$ are characterized by

$$\lim_{t \uparrow t_e} \Bigg[-\lambda \left|t - t_e\right|^{\lambda-1} \left(y_K(\omega) - \int_{t_e-\rho}^{t} \frac{\tilde{W}_s(\omega) F_{K,1}(s) + F_{K,2}(s)}{\left|s - t_e\right|^{1+\lambda}} \, ds \right) - \left(\frac{\tilde{W}_t(\omega) F_{K,1}(t) + F_{K,2}(t)}{\left|t - t_e\right|} \right) \Bigg]. \tag{6.83}$$

More precisely,

$$\begin{aligned} \{(6.83) = +\infty\} \quad &\Longleftrightarrow \quad \left\{ \lim_{t \uparrow t_e} \theta_t^{u,\star}(\omega) = [+\infty, \ldots, +\infty]^\top \right\} \\ \{(6.83) = -\infty\} \quad &\Longleftrightarrow \quad \left\{ \lim_{t \uparrow t_e} \theta_t^{u,\star}(\omega) = [-\infty, \ldots, -\infty]^\top \right\}. \end{aligned} \tag{6.84}$$

Suggestively, we first rewrite the expression in (6.83) as

$$\begin{aligned} &F_{K,2}(t) \left(\lambda \left|t - t_e\right|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{1}{\left|s - t_e\right|^{1+\lambda}} \, ds - \frac{1}{\left|t - t_e\right|} \right) \\ &\quad + \lambda \left|t - t_e\right|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{F_{K,2}(s) - F_{K,2}(t)}{\left|s - t_e\right|^{1+\lambda}} \, ds \\ &\quad - \lambda \left|t - t_e\right|^{\lambda-1} y_K(\omega) \\ &\quad + \tilde{W}_t(\omega) F_{K,1}(t) \left(\lambda \left|t - t_e\right|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{1}{\left|s - t_e\right|^{1+\lambda}} \, ds - \frac{1}{\left|t - t_e\right|} \right) \\ &\quad + \lambda \left|t - t_e\right|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{\tilde{W}_s(\omega) \left[F_{K,1}(s) - F_{K,1}(t)\right]}{\left|s - t_e\right|^{1+\lambda}} \, ds \\ &\quad + \lambda \left|t - t_e\right|^{\lambda-1} F_{K,1}(t) \int_{t_e-\rho}^{t} \frac{\tilde{W}_s(\omega) - \tilde{W}_t(\omega)}{\left|s - t_e\right|^{1+\lambda}} \, ds \end{aligned} \tag{6.85}$$

Let R_1 and R_2 be the deterministic Lipschitz coefficients for $F_{K,1}$ and $F_{K,2}$. The first two lines of (6.85) are deterministic, and we can obtain the following bounds:

$$\begin{aligned}
&\left| F_{K,2}(t) \left(\lambda |t - t_e|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{1}{|s - t_e|^{1+\lambda}}\, ds - \frac{1}{|t - t_e|} \right) \right| \\
&\qquad \leq \frac{|F_{K,2}(t)|\, |t - t_e|^{\lambda-1}}{\rho^{\lambda}} \\
&\left| \lambda |t - t_e|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{F_{K,2}(s) - F_{K,2}(t)}{|s - t_e|^{1+\lambda}}\, ds \right| \\
&\qquad \leq \left(\frac{\lambda R_2}{1 - \lambda} \right) \left(\rho^{1-\lambda} |t - t_e|^{\lambda-1} - 1 \right) \qquad (6.86)
\end{aligned}$$

In (6.85), the third line is deterministic conditional on $\tilde{\mathcal{F}}_{t_e-\rho}$. Lines 4 - 6 of (6.85) are stochastic conditional on $\tilde{\mathcal{F}}_{t_e-\rho}$. Letting $R_3(\omega)$ be the maximum of $\left|\tilde{W}_t(\omega)\right|$ on $[t_e - \rho, t_e]$, we notice that

$$\begin{aligned}
&\left| \tilde{W}_t(\omega)\, F_{K,1}(t) \left(\lambda |t - t_e|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{1}{|s - t_e|^{1+\lambda}}\, ds - \frac{1}{|t - t_e|} \right) \right| \\
&\qquad \leq \frac{F_{K,1}(t) \left|\tilde{W}_t(\omega)\right| |t - t_e|^{\lambda-1}}{\rho^{\lambda}} \\
&\left| \lambda |t - t_e|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{\tilde{W}_s(\omega) \left[F_{K,1}(s) - F_{K,1}(t)\right]}{|s - t_e|^{1+\lambda}}\, ds \right| \\
&\qquad \leq \left(\frac{\lambda R_1 R_3(\omega)}{1 - \lambda} \right) \left(\rho^{1-\lambda} |t - t_e|^{\lambda-1} - 1 \right). \qquad (6.87)
\end{aligned}$$

When $\lambda > 1$, it follows that we see that (6.83) has the same behavior as

$$\lim_{t \uparrow t_e} \left[|t - t_e|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{\tilde{W}_s(\omega) - \tilde{W}_t(\omega)}{|s - t_e|^{1+\lambda}}\, ds \right] \qquad (6.88)$$

(all other terms tend to 0 $\tilde{P}$-a.s.). Using integration by parts,

$$\begin{aligned}
&|t - t_e|^{\lambda-1} \int_{t_e-\rho}^{t} \frac{\tilde{W}_s(\omega) - \tilde{W}_t(\omega)}{|s - t_e|^{1+\lambda}}\, ds \\
&\qquad \sim \mathcal{N}\left(0, |t - t_e|^{2\lambda-2} \int_{t_e-\rho}^{t} \left(\frac{|s - t_e|^{-\lambda}}{\lambda} - \frac{1}{\lambda \rho^{\lambda}} \right)^2 ds \right). \qquad (6.89)
\end{aligned}$$

Asymptotically, the variance in (6.89) grows like $|t - t_e|^{-1}$ as $t \uparrow t_e$, completing the proof of Lemma VI.42. □

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] *Bats: routing strategies (FIX V2 routing tags and instructions)*. http://cdn.batstrading.com/resources/features/bats_exchange_routing-strategies.pdf. Accessed: 2017-05-18.

[2] *BBC News: probe into Japan share sale error*. http://news.bbc.co.uk/2/hi/business/4512962.stm. Accessed: 2017-04-15.

[3] *Commodity Futures Trading Commission and the Securities & Exchange Commission: findings regarding the market events of May 6, 2010*. https://www.sec.gov/news/studies/2010/marketevents-report.pdf. Accessed: 2017-04-15.

[4] *Commodity Futures Trading Commission: CFTC charges U.K. resident Navinder Singh Sarao and his company Nav Sarao Futures Limited PLC with price manipulation and spoofing (RELEASE: pr7156-15)*. http://www.cftc.gov/PressRoom/PressReleases/pr7156-15. Accessed: 2017-04-15.

[5] *Consolidated Tape Association: overview*. https://www.ctaplan.com/index. Accessed: 2017-04-15.

[6] *EuroMillions information: pari-mutuel*. https://www.euro-millions.com/pari-mutuel. Accessed: 2016-03-01.

[7] *International Federation of Horseracing Authorities 2014 annual report*. http://www.horseracingintfed.com/default.asp?section=Resources&area=4&FF=15&statsyear=2014&report=D. Accessed: 2016-03-01.

[8] *Nanex LLC.: flash crash summary report*. http://www.nanex.net/FlashCrashFinal/FlashCrashSummary.html. Accessed: 2017-04-15.

[9] *Nanex LLC.: HFT broke the stock market*. http://www.nanex.net/aqck2/4178.html. Accessed: 2017-04-15.

[10] *Nanex LLC.: NxResearch*. http://www.nanex.net/NxResearch/. Accessed: 2017-04-15.

[11] *National Energy Board: propane market review (2016 update)*. https://www.neb-one.gc.ca/nrg/sttstc/ntrlgslqds/rprt/2016/2016prpn-eng.html. Accessed: 2017-04-15.

[12] *Nonlocal equations wiki*. https://www.ma.utexas.edu/mediawiki/index.php/Starting_page. Accessed: 2015-01-25.

[13] *NYSE Arca: Rule 7.10 (clearly erroneous executions)*. https://www.nyse.com/publicdocs/nyse/markets/nyse-arca/NYSEArcaRule7.10.pdf. Accessed: 2017-04-15.

[14] *Russell Investments: 2015 Index Reconstitution*. http://www.russell.com/indexes/americas/tools-resources/reconstitution/default.page. Accessed: 2015-03-15.

[15] *Securities & Exchange Commission: order approving, on a pilot basis, the National Market System Plan to address extraordinary market volatility.* `https://www.sec.gov/rules/sro/nms/2012/34-67091.pdf`. Accessed: 2017-04-15.

[16] *Securities & Exchange Commission: Rule 611 of Regulation NMS.* `https://www.sec.gov/spotlight/emsac/memo-rule-611-regulation-nms.pdf`. Accessed: 2017-04-15.

[17] *Securities and Exchange Commission: Merrill Lynch charged with trading control failures that led to mini flash crashes.* `https://www.sec.gov/news/pressrelease/2016-192.html`. Accessed: 2017-04-15.

[18] B. Acciaio, M. Beiglböck, F. Penkner, and W. Schachermayer, *A model-free version of the fundamental theorem of asset pricing and the super-replication theorem*, Math. Finance, 26 (2016), pp. 233–251.

[19] R. Almgren and N. Chriss, *Value under liquidation*, Risk, 12 (1999), pp. 61–63.

[20] ——, *Optimal execution of portfolio transactions*, J. Risk, 3 (2001), pp. 5–40.

[21] R. Almgren and J. Lorenz, *Adaptive arrival price*, Trading, 2007 (2007), pp. 59–66.

[22] R. F. Almgren, *Optimal execution with nonlinear impact functions and trading-enhanced risk*, Appl. Math. Finance, 10 (2003), pp. 1–18.

[23] C. A. Anderson, M. R. Lepper, and L. Ross, *Perseverance of social theories: the role of explanation in the persistence of discredited information*, J. Pers. Soc. Psychol., 39 (1980), p. 1037.

[24] W. B. Arthur, *Inductive reasoning and bounded rationality*, Am. Econ. Rev., 84 (1994), pp. 406–411.

[25] P. Asch, B. G. Malkiel, and R. E. Quandt, *Racetrack betting and informed behavior*, J. Financ. Econ., 10 (1982), pp. 187 – 194.

[26] M. Attari, A. S. Mello, and M. E. Ruckes, *Arbitraging arbitrageurs*, J. Finance, 60 (2005), pp. 2471–2511.

[27] R. J. Aumann, *Markets with a continuum of traders*, Econometrica, 32 (1964), pp. 39–50.

[28] M. Avellaneda, A. Levy, and A. Parás, *Pricing and hedging derivative securities in markets with uncertain volatilities*, Appl. Math. Finance, 2 (1995), pp. 73–88.

[29] J. Azar, S. Raina, and M. C. Schmalz, *Ultimate ownership and bank competition.* `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2710252`. Accessed: 2017-05-20.

[30] J. Azar, M. C. Schmalz, and I. Tecu, *Anti-competitive effects of common ownership.* `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2427345`. Accessed: 2017-05-20.

[31] T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory, 2nd Edition*, Society for Industrial and Applied Mathematics, 1998.

[32] E. J. Balder, *On the existence of Cournot-Nash equilibria in continuum games*, J. Math. Econom., 32 (1999), pp. 207 – 223.

[33] ——, *A unifying pair of Cournot-Nash equilibrium existence results*, J. Econom. Theory, 102 (2002), pp. 437 – 470.

[34] P. Bank, H. M. Soner, and M. Voss, *Hedging with temporary price impact*, Math. Financ. Econ., 11 (2017), pp. 215–239.

[35] N. Barberis, A. Shleifer, and R. Vishny, *A model of investor sentiment*, J. Financ. Econ., 49 (1998), pp. 307 – 343.

[36] V. Barbu, *Differential equations*, Springer Undergraduate Mathematics Series, Springer, Cham, 2016. Translated from the 1985 Romanian original by Liviu Nicolaescu.

[37] G. I. Barenblatt, *Similarity, self-similarity, and intermediate asymptotics*, Consultants Bureau [Plenum], New York-London, 1979. Translated from the Russian by N. Stein, translation edited by M. Van Dyke, with a foreword by Y. B. Zel'dovich.

[38] K. Baron and J. Lange, *Parimutuel applications in finance*, Palgrave Macmillan UK, 2007.

[39] R. Bass and D. Levin, *Harnack inequalities for jump processes*, Potential Anal., 17 (2002), pp. 375–388.

[40] E. Bayraktar, A. Cosso, and H. Pham, *Robust feedback switching control: dynamic programming and viscosity solutions*, SIAM J. Control Optim., 54 (2016), pp. 2594–2628.

[41] E. Bayraktar and Y.-J. Huang, *Robust maximization of asymptotic growth under covariance uncertainty*, Ann. Appl. Probab., 23 (2013), pp. 1817–1840.

[42] E. Bayraktar, Y.-J. Huang, and Z. Zhou, *On hedging American options under model uncertainty*, SIAM J. Financial Math., 6 (2015), pp. 425–447.

[43] E. Bayraktar, I. Karatzas, and S. Yao, *Optimal stopping for dynamic convex risk measures*, Illinois J. Math., 54 (2010), pp. 1025–1067 (2012).

[44] E. Bayraktar and M. Ludkovski, *Optimal trade execution in illiquid markets*, Math. Finance, 21 (2011), pp. 681–701.

[45] ——, *Liquidation in limit order books with controlled intensity*, Math. Finance, 24 (2014), pp. 627–650.

[46] E. Bayraktar and A. Munk, *Mini-flash crashes, model risk, and optimal execution.* `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2975769`. Accessed: 2017-06-02.

[47] ——, *Comparing the G-normal distribution to its classical counterpart*, Commun. Stoch. Anal., 9 (2015), pp. 1–18.

[48] ——, *An α-stable limit theorem under sublinear expectation*, Bernoulli, 22 (2016), pp. 2548–2578.

[49] ——, *Are the high-rolling quants of horse racing our friends or foes?*, The Conversation, (2016).

[50] ——, *High-roller impact: A large generalized game model of parimutuel wagering*, Market Microstructure and Liquidity, (forthcoming).

[51] E. Bayraktar and Y. Zhang, *Minimizing the probability of lifetime ruin under ambiguity aversion*, SIAM J. Control Optim., 53 (2015), pp. 58–90.

[52] ——, *Fundamental theorem of asset pricing under transaction costs and model uncertainty*, Math. Oper. Res., 41 (2016), pp. 1039–1054.

[53] M. Beiglböck, P. Henry-Labordère, and F. Penkner, *Model-independent bounds for option prices—a mass transport approach*, Finance Stoch., 17 (2013), pp. 477–501.

[54] D. Bertsimas and A. W. Lo, *Optimal control of execution costs*, J. Financ. Mark., 1 (1998), pp. 1–50.

[55] L. E. Blume and D. Easley, *Learning to be rational*, J. Econom. Theory, 26 (1982), pp. 340–351.

[56] R. N. Bolton and R. G. Chapman, *Searching for positive returns at the track: a multinomial logit model for handicapping horse races*, Management Sci., 32 (1986), pp. 1040–1060.

[57] É. Borel, *Sur le pari mutuel*, Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, 207 (1938), pp. 197–200.

[58] B. Bouchard and M. Nutz, *Arbitrage and duality in nondominated discrete-time models*, Ann. Appl. Probab., 25 (2015), pp. 823–859.

[59] G. E. P. Box, *Science and statistics*, J. Amer. Statist. Assoc., 71 (1976), pp. 791–799.

[60] M. Bray, *Learning, estimation, and the stability of rational expectations*, J. Econom. Theory, 26 (1982), pp. 318–339.

[61] J. Brogaard, M. Ringgenberg, and D. Sovich, *The economic impact of index investing.* `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2663398`. Accessed: 2017-05-20.

[62] M. K. Brunnermeier and L. H. Pedersen, *Predatory trading*, The Journal of Finance, 60 (2005), pp. 1825–1863.

[63] W. Buffett, *Berkshire Hathaway, Inc.: 1996 Chairman's Letter.* `http://www.berkshirehathaway.com/letters/1996.html`. Accessed: 2015-03-15.

[64] L. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*, vol. 43 of American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 1995.

[65] L. Caffarelli and L. Silvestre, *Regularity theory for fully nonlinear integro-differential equations*, Comm. Pure Appl. Math., 62 (2009), pp. 597–638.

[66] ——, *The Evans-Krylov theorem for nonlocal fully nonlinear equations*, Ann. of Math. (2), 174 (2011), pp. 1163–1187.

[67] J. Cai and T. Houge, *Long-term impact of Russell 2000 Index rebalancing*, Financ. Anal. J., 64 (2008), pp. 76–91.

[68] F. Canyameres, *L'Homme de la Belle Époque*, les Éditions universelles, Paris, 1946.

[69] B. I. Carlin, M. S. Lobo, and S. Viswanathan, *Episodic liquidity crises: cooperative and predatory trading*, J. Finance, 62 (2007), pp. 2235–2274.

[70] G. Carmona and K. Podczeck, *Existence of Nash equilibrium in games with a measure space of players and discontinuous payoff functions*, J. Econom. Theory, 152 (2014), pp. 130 – 178.

[71] R. Carmona and D. Lacker, *A probabilistic weak formulation of mean field games and applications*, Ann. Appl. Probab., 25 (2015), pp. 1189–1231.

[72] R. A. Carmona and J. Yang, *Predatory trading: a game on volatility and liquidity.* `http://www.princeton.edu/rcarmona/download/fe/PredatoryTradingGameQF.pdf`. Accessed: 2015-03-15.

[73] A. Cartea, R. Donnelly, and S. Jaimungal, *Portfolio liquidation and ambiguity aversion.* `https://papers-ssrn-com.proxy.lib.umich.edu/sol3/papers.cfm?abstract_id=2946136`. Accessed: 2017-05-11.

[74] ——, *Algorithmic trading with model uncertainty*, SIAM J. Financial Math., (forthcoming).

[75] Á. Cartea, S. Jaimungal, and D. Kinzebulatov, *Algorithmic trading with learning*, Int. J. Theoretical Appl. Finance, 19 (2016), p. 1650028.

[76] A. Cartea, S. Jaimungal, and Z. Qin, *Model uncertainty in commodity markets*, SIAM J. Financial Math., 7 (2016), pp. 1–33.

[77] S. Chadha and R. E. Quandt, *Betting bias and market equilibrium in racetrack betting*, Appl. Finan. Econ., 6 (1996), pp. 287–292.

[78] S. Chakravarty, P. Jain, J. Upson, and R. Wood, *Clean sweep: informed trading through intermarket sweep orders*, J. Financ. Quant. Anal., 47 (2012), pp. 414–435.

[79] A. Chalmers, *What is this thing called science?*, Hackett Publishing Company and the University of Queensland Press, 2013.

[80] P. Chan and R. Sircar, *Bertrand and Cournot mean field games*, Appl. Math. Optim., 71 (2015), pp. 533–569.

[81] ——, *Fracking, renewables & mean field games*, SIAM Rev., (forthcoming).

[82] H. Chen, G. Noronha, and V. Singal, *The price response to S&P 500 Index additions and deletions: evidence of asymmetry and a new explanation*, J. Finance, 59 (2004), pp. 1901–1930.

[83] ——, *Index changes and losses to index fund investors*, Financ. Anal. J., 62 (2006), pp. 31–47.

[84] H.-L. Chen, *On Russell index reconstitution*, Review of Quantitative Finance and Accounting, 26 (2006), pp. 409–430.

[85] R. Cheung, *Pari-mutuel wagering.* Undergraduate Thesis, Princeton University, 1985.

[86] C. S. Chu, A. Lehnert, and W. Passmore, *Strategic trading in multiple assets and the effects on market volatility*, Int. J. Cent. Bank., 5 (2009), pp. 143–172.

[87] E. A. Coddington and R. Carlson, *Linear ordinary differential equations*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[88] K. Colaneri, Z. Eksi, R. Frey, and M. Szölgyenyi, *Shall I sell or shall I wait? Optimal liquidation under partial information with price impact*, ArXiv e-prints, (2016).

[89] R. Cont, *Model uncertainty and its impact on the pricing of derivative instruments*, Math. Finance, 16 (2006), pp. 519–547.

[90] S. Correa and J. P. Torres-Martínez, *Essential equilibria of large generalized games*, Econom. Theory, 57 (2014), pp. 479–513.

[91] L. Denis and C. Martini, *A theoretical framework for the pricing of contingent claims in the presence of model uncertainty*, Ann. Appl. Probab., 16 (2006), pp. 827–852.

[92] D. Dick, *Erroneous combustion*, CFA Institute Magazine, 24 (2013), pp. 20–21.

[93] Y. Dolinsky, *Numerical schemes for G-expectations*, Electron. J. Probab., 17 (2012), pp. 1–15.

[94] Y. Dolinsky, M. Nutz, and H. M. Soner, *Weak approximation of G-expectations*, Stochastic Process. Appl., 122 (2012), pp. 664–675.

[95] Y. Dolinsky and H. M. Soner, *Martingale optimal transport and robust hedging in continuous time*, Probab. Theory Related Fields, 160 (2014), pp. 391–427.

[96] D. Easley, M. M. López de Prado, and M. OHara, *The microstructure of the "flash crash": flow toxicity, liquidity crashes, and the probability of informed trading*, J. Portfolio Manage., 37 (2011), pp. 118–128.

[97] D. Easley, M. M. López de Prado, and M. O'Hara, *Flow toxicity and liquidity in a high-frequency world*, Rev. Financ. Stud., 25 (2012), p. 1457.

[98] E. Ekström and J. Vaicenavicius, *Optimal liquidation of an asset under drift uncertainty*, SIAM J. Financial Math., 7 (2016), pp. 357–381.

[99] E. Elhauge, *Horizontal shareholding.* https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2632024. Accessed: 2017-05-20.

[100] I. Esponda, *Behavioral equilibrium in economies with adverse selection*, Am. Econ. Rev., 98 (2008), pp. 1269–1291.

[101] G. W. Evans and S. Honkapohja, *Learning and expectations in macroeconomics*, Princeton University Press, 2001.

[102] M. Farrell, *Mini flash crashes: a dozen a day.* http://money.cnn.com/2013/03/20/investing/mini-flash-crash/. Accessed: 2017-04-15.

[103] R. Feeney and S. P. King, *Sequential parimutuel games*, Econom. Lett., 72 (2001), pp. 165 – 173.

[104] R. A. Ferri and A. C. Benke, *A case for index fund portfolios.* http://www.rickferri.com/WhitePaper.pdf. Accessed: 2015-03-15.

[105] Fidelity, *Index funds.* https://www.fidelity.com/mutual-funds/index-funds/overview. Accessed: 2015-03-15.

[106] V. Filimonov and D. Sornette, *Quantifying reflexivity in financial markets: Toward a prediction of flash crashes*, Phys. Rev. E., 85 (2012), p. 056108.

[107] P. A. Forsyth, *A Hamilton-Jacobi-Bellman approach to optimal trade execution*, Appl. Numer. Math., 61 (2011), pp. 241–265.

[108] P. A. Forsyth, J. S. Kennedy, S. T. Tse, and H. Windcliff, *Optimal trade execution: a mean quadratic variation approach*, J. Econom. Dynam. Control, 36 (2012), pp. 1971–1991.

[109] R. Frey, *Superreplication in stochastic volatility models and optimal stopping*, Finance Stoch., 4 (2000), pp. 161–187.

[110] R. Frey, A. Gabih, and R. Wunderlich, *Portfolio optimization under partial information with expert opinions*, Int. J. Theor. Appl. Finance, 15 (2012), p. 1250009.

[111] D. Fudenberg and D. K. Levine, *Self-confirming equilibrium*, Econometrica, 61 (1993), pp. 523–545.

[112] S. G. Shankar and J. M. Miller, *Market reaction to changes in the S&P SmallCap 600 Index*, Finan. Rev., 41 (2006), pp. 339–360.

[113] N. Gârleanu and L. H. Pedersen, *Dynamic trading with predictable returns and transaction costs*, J. Finance, 68 (2013), pp. 2309–2340.

[114] J. Gatheral and A. Schied, *Optimal trade execution under geometric Brownian motion in the Almgren and Chriss framework*, Int. J. Theor. Appl. Finance, 14 (2011), pp. 353–368.

[115] R. Gayduk and S. Nadtochiy, *Liquidity Effects of Trading Frequency*, Math. Finance, (forthcoming).

[116] A. Golub, J. Keane, and S.-H. Poon, *High frequency trading and mini flash crashes*, ArXiv e-prints, (2012).

[117] D. A. Gomes and V. K. Voskanyan, *Extended deterministic mean-field games*, ArXiv e-prints, (2013).

[118] S. E. Gorman and J. M. Gorman, *Denying to the grave: why we ignore the facts that will save us*, Oxford University Press, 2016.

[119] T. C. Green and R. Jame, *Strategic trading by index funds and liquidity provision around S&P 500 Index additions*, J. Financ. Mark., 14 (2011), pp. 605–624.

[120] L. P. Hansen and T. J. Sargent, *Robustness*, Princeton University Press, Princeton, NJ, 2008.

[121] D. B. Hausch, V. S. Lo, and W. T. Ziemba, eds., *Efficiency of Racetrack Betting Markets*, World Scientific, Singapore, 2008 ed.

[122] D. B. Hausch and W. T. Ziemba, eds., *Handbook of sports and lottery markets*, Handbooks in Finance, Elsevier, San Diego, 2008.

[123] D. B. Hausch, W. T. Ziemba, and M. Rubinstein, *Efficiency of the market for racetrack betting*, Management Sci., 27 (1981), pp. 1435–1452.

[124] S. Herrmann, J. Muhle-Karbe, and F. T. Seifried, *Hedging with small uncertainty aversion*, Finance Stoch., 21 (2017), pp. 1–64.

[125] D. Hobson, *The Skorokhod embedding problem and model-independent bounds for option prices*, in Paris-Princeton Lectures on Mathematical Finance 2010, vol. 2003 of Lecture Notes in Math., Springer, Berlin, 2011, pp. 267–318.

[126] D. G. Hobson, *Volatility misspecification, option pricing and superreplication via coupling*, Ann. Appl. Probab., 8 (1998), pp. 193–205.

[127] M. Hu, S. Ji, S. Peng, and Y. Song, *Comparison theorem, Feynman-Kac formula and Girsanov transformation for BSDEs driven by G-Brownian motion*, Stochastic Process. Appl., 124 (2014), pp. 1170–1195.

[128] M. Hu and X. Li, *Independence under the G-expectation framework*, J. Theoret. Probab., forthcoming (2012).

[129] M. Hu and S. Peng, *G-Lévy processes under sublinear expectations*, ArXiv e-prints, (2009).

[130] Z.-C. Hu and L. Zhou, *Multi-dimensional central limit theorems and laws of large numbers under sublinear expectations*, Acta Math. Sin. (Engl. Ser.), (forthcoming).

[131] J. Huang and J. Wang, *Liquidity and market crashes*, Rev. Financ. Stud, 22 (2009), pp. 2607–2643.

[132] M. Huang, *Large-population LQG games involving a major player: the Nash certainty equivalence principle*, SIAM J. Control Optim., 48 (2010), pp. 3318–3353.

[133] M. Huang, R. P. Malham, and P. E. Caines, *Large population stochastic dynamic games: closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst., 6 (2006), pp. 221–252.

[134] W. Hurley and L. McDonough, *A note on the Hayek hypothesis and the favorite-longshot bias in parimutuel betting*, Am. Econ. Rev., (1995), pp. 949–955.

[135] A. Ibragimov, *G-expectations in infinite dimensional spaces and related PDEs*, ArXiv e-prints, (2013).

[136] I. A. Ibragimov and Y. V. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing, Groningen, 1971. Translated from the Russian and edited by J. F. C. Kingman.

[137] Y. Ilyashenko and S. Yakovenko, *Lectures on analytic differential equations*, vol. 86 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2008.

[138] R. Isaacs, *Optimal horse race bets*, Amer. Math. Monthly, 60 (1953), pp. 310–315.

[139] N. I. Ismail and L. Mnyanda, *Flash crash of the pound baffles traders with algorithms being blamed*. https://www.bloomberg.com/news/articles/2016-10-06/pound-plunges-6-1-percent-in-biggest-drop-since-brexit-result. Accessed: 2017-04-15.

[140] S. J. Leal, M. Napoletano, A. Roventini, and G. Fagiolo, *Rock around the clock: An agent-based model of low- and high-frequency trading*, J. Evol. Econ., 26 (2016), pp. 49–76.

[141] S. Jaimungal and M. Nourian, *Mean-field game strategies for a major-minor agent optimal execution problem*, Available at SSRN 2578733, (2015).

[142] N. Johnson, G. Zhao, E. Hunsader, H. Qi, N. Johnson, J. Meng, and B. Tivnan, *Abrupt rise of new machine ecology beyond human response time*, Sci. Rep., 3 (2013), pp. 1–7.

[143] A. Joulin, A. Lefevre, D. Grunberg, and J.-P. Bouchaud, *Stock price jumps: news and volume play a minor role*, Wilmott Magazine, Sep/Oct (2008), pp. 1–7.

[144] J. Kallsen and J. Muhle-Karbe, *High-resilience limits of block-shaped order books*, ArXiv e-prints, (2014).

[145] M. Kaplan, *The high tech trifecta*. http://www.wired.com/2002/03/betting/. Accessed: 2016-04-15.

[146] M. Kassmann and R. Schwab, *Regularity results for nonlocal parabolic equations*, Riv. Math. Univ. Parma (N.S.), 5 (2014), pp. 183–212.

[147] J. Keller, *A fake AP tweet sinks the Dow for an instant*. https://www.bloomberg.com/news/articles/2013-04-23/a-fake-ap-tweet-sinks-the-dow-for-an-instant. Accessed: 2017-04-15.

[148] A. Kim and G. Oikonomou, *Understanding index front running*. http://www.thetradenews.com/magazine/The_TRADE_Magazine/2007/December/Understanding_index_front_running.aspx. Accessed: 2015-03-15.

[149] A. A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tuzun, *The flash crash: high frequency trading in an electronic market*, J. Finance, (forthcoming).

[150] F. H. Knight, *Risk, Uncertainty, and Profit*, Houghton Mifflin Co., 1921.

[151] F. Koessler, C. Noussair, and A. Ziegelmeyer, *Parimutuel betting under asymmetric information*, J. Math. Econom., 44 (2008), pp. 733 – 744. Special Issue in Economic Theory in honor of Charalambos D. Aliprantis.

[152] F. Koessler, A. Ziegelmeyer, and M.-H. Broihanne, *The favorite-longshot bias in sequential parimutuel betting with non-expected utility players*, Theory Dec., 54 (2003), pp. 231–248.

[153] D. Kriventsov, *$C^{1,\alpha}$ interior regularity for nonlinear nonlocal elliptic equations with rough kernels*, Comm. Partial Differential Equations, 38 (2013), pp. 2081–2106.

[154] J. Lange and N. Economides, *A parimutuel market microstructure for contingent claims*, Eur. Financ. Manag., 11 (2005), pp. 25–49.

[155] H. Lara and G. Dávila, *Regularity for solutions of nonlocal, nonsymmetric equations*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 29 (2012), pp. 833–859.

[156] H. Lara and G. Dávila, *$C^{\sigma+\alpha}$ estimates for concave, non-local parabolic equations with critical drift*, ArXiv e-prints, (2014).

[157] H. Lara and G. Dávila, *Hölder estimates for non-local parabolic equations with critical drift*, ArXiv e-prints, (2014).

[158] J.-M. Lasry and P.-L. Lions, *Mean field games*, Jpn. J. Math., 2 (2007), pp. 229–260.

[159] M. Li and Y. Shi, *A general central limit theorem under sublinear expectations*, Sci. China Math., 53 (2010), pp. 1989–1994.

[160] Q. Lin, *Representation of G-martingales as stochastic integrals with respect to the G-Brownian motion*, ArXiv e-prints, (2010).

[161] ——, *General martingale characterization of G-Brownian motion*, Stoch. Anal. Appl., 31 (2013), pp. 1024–1048.

[162] R. S. Liptser and A. N. Shiryaev, *Statistics of random processes. I*, vol. 5 of Applications of Mathematics (New York), Springer-Verlag, Berlin, expanded ed., 2001. General theory, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability.

[163] P. Luo and G. Jia, *A note on characterizations of G-normal distribution*, ArXiv e-prints, (2014).

[164] A. W. Lynch and R. R. Mendenhall, *New evidence on stock price effects associated with changes in the S&P 500 Index*, J. Bus., 70 (1997), pp. 351–83.

[165] T. J. Lyons, *Uncertain volatility and the risk-free synthesis of derivatives*, Appl. Math. Finance, 2 (1995), pp. 117–133.

[166] A. Madhavan, *The Russell reconstitution effect*, Financ. Anal. J., 59 (2003), pp. 51–64.

[167] S. Mamudi, *Sudden stock crashes usually caused by human error, SEC says.* `https://www.bloomberg.com/news/articles/2013-06-18/sudden-stock-crashes-mostly-show-human-error-sec-s-berman-says`. Accessed: 2017-04-15.

[168] A. Mas-Colell, *On a theorem of Schmeidler*, J. Math. Econom., 13 (1984), pp. 201 – 206.

[169] H. Mercier and D. Sperber, *The enigma of reason*, Harvard University Press, 2017.

[170] C. C. Moallemi, B. Park, and B. Van Roy, *Strategic execution in the presence of an uninformed arbitrageur*, Journal of Financial Markets, 15 (2012), pp. 361 – 391.

[171] R. K. Narang, *Inside the Black Box*, John Wiley & Sons, Inc., 2009.

[172] A. Neufeld and M. Nutz, *Superreplication under volatility uncertainty for measurable claims*, Electron. J. Probab., 18 (2013), pp. no. 48, 14.

[173] A. Neufeld and M. Nutz, *Nonlinear Lévy processes and their characteristics*, ArXiv e-prints, (2014).

[174] S. L. Nguyen and M. Huang, *Linear-Quadratic-Gaussian mixed games with continuum-parametrized minor players*, SIAM J. Control Optim., 50 (2012), pp. 2907–2937.

[175] M. Nourian and P. E. Caines, *ϵ-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents*, SIAM J. Control Optim., 51 (2013), pp. 3302–3331.

[176] M. Nutz, *Random G-expectations*, Ann. Appl. Probab., 23 (2013), pp. 1755–1777.

[177] M. Nutz and R. van Handel, *Constructing sublinear expectations on path space*, Stochastic Process. Appl., 123 (2013), pp. 3100–3121.

[178] A. A. Obizhaeva and J. Wang, *Optimal trading strategy and supply/demand dynamics*, J. Financ. Mark., 16 (2013), pp. 1–32.

[179] K. Okada, N. Isagawa, and K. Fujiwara, *Addition to the Nikkei 225 Index and Japanese market response: temporary demand effect of index arbitrageurs*, Pac-Basin Financ. J., 14 (2006), pp. 395–409.

[180] Z. M. Onayev and V. M. Zdorovtsov, *Predatory trading around Russell reconstitution*, Available at SSRN 1101341, (2008).

[181] E. Osuka, *Girsanov's formula for G-Brownian motion*, Stochastic Process. Appl., 123 (2013), pp. 1301–1318.

[182] M. Ottaviani and P. N. Sørensen, *The timing of parimutuel bets*, `http://faculty.london.edu/mottaviani/tobaflb.pdf`. Accessed: 2016-04-15.

[183] M. Ottaviani and P. N. Sørensen, *The favorite-longshot bias: an overview of the main explanations*, in In Handbook of Sports and Lottery Markets, Elsevier North Holland, 2008.

[184] M. Ottaviani and P. N. Sørensen, *Surprised by the parimutuel odds?*, Am. Econ. Rev., 99 (2009), pp. 2129–2134.

[185] K. Paczka, *Itô calculus and jump diffusions for G-Lévy processes*, ArXiv e-prints, (2012).

[186] ——, *G-martingale representation in the G-Lévy setting*, ArXiv e-prints, (2014).

[187] F. Passerini and S. E. Vazquez, *Optimal trading with alpha predictors*, ArXiv e-prints, (2015).

[188] S. Peng, *G-Brownian motion and dynamic risk measure under volatility uncertainty*, ArXiv e-prints, (2007).

[189] ——, *G-expectation, G-Brownian motion and related stochastic calculus of Itô type*, in Stochastic Analysis and Applications, vol. 2 of Abel Symp., Springer, Berlin, 2007, pp. 541–567.

[190] ——, *Law of large numbers and central limit theorem under nonlinear expectations*, ArXiv e-prints, (2007).

[191] ——, *A new central limit theorem under sublinear expectations*, ArXiv e-prints, (2008).

[192] ——, *Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation*, Stochastic Process. Appl., 118 (2008), pp. 2223–2253.

[193] ——, *Survey on normal distributions, central limit theorem, Brownian motion and the related stochastic calculus under sublinear expectations*, Sci. China Ser. A, 52 (2009), pp. 1391–1411.

[194] ——, *Nonlinear expectations and stochastic calculus under uncertainty*, ArXiv e-prints, (2010).

[195] S. Peng, Y. Song, and J. Zhang, *A complete representation theorem for G-martingales*, Stochastics, 86 (2014), pp. 609–631.

[196] D. M. Pennock, *A dynamic pari-mutuel market for hedging, wagering, and information aggregation*, in Proceedings of the 5th ACM Conference on Electronic Commerce, EC '04, New York, NY, USA, 2004, ACM, pp. 170–179.

[197] M. Peters, A. M.-C. So, and Y. Ye, *Internet and Network Economics: Third International Workshop, WINE 2007, San Diego, CA, USA, December 12-14, 2007. Proceedings*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, ch. Pari-Mutuel Markets: Mechanisms and Performance, pp. 82–95.

[198] M. Phillips, *How software updates are destroying the stock market*. https://www.bloomberg.com/news/articles/2012-08-03/how-software-updates-are-destroying-the-stock-market. Accessed: 2017-04-15.

[199] B. Pisani, *What happened to Kraft today?* http://www.cnbc.com/id/49279166. Accessed: 2017-04-15.

[200] R. Piziak and P. L. Odell, *Full rank factorization of matrices*, Math. Mag., 72 (1999), pp. 193–201.

[201] R. C. Plott, J. Wit, and C. W. Yang, *Parimutuel betting markets as information aggregation devices: experimental results*, Econom. Theory, 22, pp. 311–351.

[202] A. D. Polyanin and A. V. Manzhirov, *Handbook of integral equations*, CRC press, 2012.

[203] D. Possamaï, G. Royer, and N. Touzi, *On the robust superhedging of measurable claims*, Electron. Commun. Probab., 18 (2013), pp. no. 95, 13.

[204] J. Potters and J. Wit, *Bets and bids: favorite-longshot bias and winner's curse*, CentER Discussion Paper, 1996-04.

[205] R. E. Quandt, *Betting and equilibrium*, Quart. J. Econom., 101 (1986), pp. 201–207.

[206] A. R. Villegas and J. P. Torres-Martínez, *On pure strategy equilibria in large generalized games*. https://mpra.ub.uni-muenchen.de/46840/. Accessed: 2016-04-15.

[207] K. P. Rath, *A direct proof of the existence of pure strategy equilibria in games with a continuum of players*, Econom. Theory, 2 (1992), pp. 427–433.

[208] L. Ren, *On representation theorem of sublinear expectation related to G-Lévy process and paths of G-Lévy process*, Statist. Probab. Lett., 83 (2013), pp. 1301–1310.

[209] B. Rosner, *Optimal allocation of resources in a pari-mutuel setting*, Management Sci., 21 (1975), pp. 997–1006.

[210] T. J. Sargent, *Bounded rationality in macroeconomics: The Arne Ryde memorial lectures*, Oxford University Press, 1993.

[211] J. A. Scheinkman and W. Xiong, *Overconfidence and speculative bubbles*, J. Polit. Econ., 111 (2003), pp. 1183–1220.

[212] A. Schied, *Robust strategies for optimal order execution in the Almgren-Chriss framework*, Appl. Math. Finance, 20 (2013), pp. 264–286.

[213] A. Schied and T. Zhang, *A state-constrained differential game arising in optimal portfolio liquidation*, ArXiv e-prints, (2013).

[214] D. Schmeidler, *Equilibrium points of nonatomic games*, J. Stat. Phys., 7 (1973), pp. 295–300.

[215] T. Schöneborn, *Trade execution in illiquid markets: optimal stochastic control and multi-agent equilibria*, PhD thesis, Universitätsbibliothek, 2008.

[216] T. Schöneborn and A. Schied, *Liquidation in the face of adversity: stealth vs. sunshine trading*, in EFA 2008 Athens Meetings Paper, 2009.

[217] R. Schwab, M. Rang, and M. Kassmann, *Integro-differential equations with nonlinear directional dependence*, Indiana Univ. Math. J., 63 (2014), pp. 1467–1498.

[218] R. Schwab and L. Silvestre, *Regularity for parabolic integro-differential equations with very irregular kernels*, ArXiv e-prints, (2014).

[219] J. Serra, *$C^{\sigma+\alpha}$ regularity for concave nonlocal fully nonlinear elliptic equations with rough kernels*, ArXiv e-prints, (2014).

[220] L. Silvestre, *On the differentiability of the solution to the Hamilton-Jacobi equation with critical fractional diffusion*, Adv. Math., 226 (2011), pp. 2020 – 2039.

[221] L. Silvestre, *Regularity estimates for parabolic integro-differential equations and applications*, Proceedings of the International Congress of Mathematicians, (forthcoming).

[222] S. Sloman and P. Fernbach, *The knowledge illusion: why we never think alone*, Penguin, 2017.

[223] H. M. Soner, N. Touzi, and J. Zhang, *Martingale representation theorem for the G-expectation*, Stochastic Process. Appl., 121 (2011), pp. 265–287.

[224] Y. Song, *Some properties on G-evaluation and its applications to G-martingale decomposition*, Sci. China Math., 54 (2011), pp. 287–300.

[225] ——, *Characterizations of processes with stationary and independent increments under G-expectation*, Ann. Inst. Henri Poincaré Probab. Stat., 49 (2013), pp. 252–269.

[226] N. N. Taleb, *The black swan : the impact of the highly improbable*, Random House, 2 ed., 2010.

[227] D. Terrell and A. Farmer, *Optimal betting and efficiency in parimutuel betting markets with information costs*, Econ. J., 106 (1996), pp. 846–868.

[228] R. H. Thaler and W. T. Ziemba, *Parimutuel betting markets: racetracks and lotteries*, J. Econ. Perspect., 2 (1988), pp. 161–74.

[229] R. M. Thrall, *Some results in non-linear programming*, in Proceedings of the Second Symposium in Linear Programming, Washington, D. C., 1955, National Bureau of Standards, Washington, D. C., 1955, pp. 471–493.

[230] Vanguard, *Index funds could help lower long-term costs*. `https://investor.vanguard.com/mutual-funds/index-funds`. Accessed: 2013-03-15.

[231] W. Walter, *Ordinary differential equations*, vol. 182 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1998. Translated from the sixth German (1996) edition by Russell Thompson, Readings in Mathematics.

[232] Y. Wang, F. Yu, H. Tang, and M. Huang, *A mean field game theoretic approach for security enhancements in mobile ad hoc networks*, IEEE T. Wirel. Commun., 13 (2014), pp. 1616–1627.

[233] W. Wasow, *Asymptotic expansions for ordinary differential equations*, Pure and Applied Mathematics, Vol. XIV, Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney, 1965.

[234] T. Watanabe, *A parimutuel system with two horses and a continuum of bettors*, J. Math. Econom., 28 (1997), pp. 85 – 100.

[235] T. Watanabe, H. Nonoyama, and M. Mori, *A model of a general parimutuel system: characterizations and equilibrium selection*, Internat. J. Game Theory, 23, pp. 237–260.

[236] J. Wathen, *A billionaire's warning on index funds.* http://money.cnn.com/2015/03/31/investing/investing-index-funds-warning/. Accessed: 2015-03-15.

[237] R. J. Weber, *Noncooperative games*, in Game theory and its applications (Biloxi, Miss., 1979), vol. 24 of Proc. Sympos. Appl. Math., Amer. Math. Soc., Providence, R.I., 1981, pp. 83–125.

[238] A. Wiszniewska-Matyszkiel, *Dynamic oligopoly as a mixed large game–toy market*, in Mathematical programming and game theory for decision making, vol. 1, World Scientific, Singapore, 2008, pp. 369–390.

[239] M. Wyart, J.-P. Bouchaud, J. Kockelkoren, M. Potters, and M. Vettorazzo, *Relation between bid-ask spread, impact and volatility in order-driven markets*, Quantitative Finance, 8 (2008), pp. 41–57.

[240] C. S. Xu, *Predatory trading on Russell Index rebalance dates: flow abnormalities and a case for front-running.* Undergraduate Thesis, Princeton University, 2012.

[241] J. Xu, H. Shang, and B. Zhang, *A Girsanov type theorem under G-framework*, Stoch. Anal. Appl., 29 (2011), pp. 386–406.

[242] J. Xu and B. Zhang, *Martingale characterization of G-Brownian motion*, Stochastic Process. Appl., 119 (2009), pp. 232–248.

[243] ——, *Martingale property and capacity under G-framework*, Electron. J. Probab., 15 (2010), pp. 2041–2068.

[244] D. Zhang and Z. Chen, *A weighted central limit theorem under sublinear expectations*, Comm. Statist. Theory Methods, 43 (2014), pp. 566–577.